Cultural distance, gender and praise in peer review

Guangyao Zhang, Lili Wang & Xianwen Wang

Published online: 03 Oct 2024.

Submit your article to this journal ⬚

View related articles ⬚

View Crossmark data ⬚

Taylor & Francis
Taylor & Francis Group

Check for updates

Cultural distance, gender and praise in peer review

Guangyao Zhang[a,b,c], Lili Wang[c], and Xianwen Wang[a,b]

[a]School of Public Administration and Policy, Dalian University of Technology, Dalian, China; [b]WISE Lab, Institute of Science of Science and S&T Management, Dalian University of Technology, Dalian, China; [c]UNU-MERIT, Maastricht University, Maastricht, The Netherlands

ABSTRACT

Background: Understanding review comments holds significant importance within the realm of scientific discourse. This study aims to conduct an empirical analysis of factors associated with praise in peer review.
Methods: The study involved manual labeling of "praise" in 952 review comments drawn from 301 articles published in the British Medical Journal, followed by regression analysis.
Results: The study reveals that authors tend to receive longer praise when they share a cultural proximity with the reviewers. Additionally, it is observed that female reviewers are more inclined to provide praise
Conclusions: In summary, these discoveries contribute valuable insights for the development of a constructive peer review process and the establishment of a more inclusive research culture.

1. Introduction

Peer review is the cornerstone of scientific publishing, playing an important role in the evolution of the scientific community and the progression of research endeavors (Crijns, Ottenhoff, andwct Ring 2021). However, the peer review process is at times marked by a perceived rigor, bordering on occasional harshness, if not outright severity. To uphold the standards of scholarly publications and maintain the integrity of scientific inquiry, reviewers find themselves compelled to adopt a stringent attitude, offering pointed critiques. A investigation of manuscripts submitted to the *American Sociological Review*, as conducted by Bakanic, McPhail, and Simon (1987), revealed a prevalence of negative comments compared to positive ones, aligning with the principles of critical rationalism articulated by Popper (1961) that advocate for rational criticism to propel scholarly advancements.

The issuance of robust criticism undoubtedly poses a significant challenge for authors (Bharti et al. 2023). The mean tone of review comments, as a potential source of bias, places researchers, particularly those in the

---

early stages of their careers, at a disadvantage (Baglini and Parsons 2020; Clements 2020; Parsons and Baglini 2021). Given the highly competitive nature of the scientific landscape, the associated pressure to succeed tends to amplify frustrations among researchers. This is especially palpable in the context of peer review, acknowledged as one of the most rigorous processes in scientific inquiry, where journal decisions occasionally prove disheartening. Consequently, there is a growing recognition within the academic sphere of the need for "kindness" in research culture (Derrick 2020, 2021).

Early journal editors have traditionally emphasized the role of reviewers in quality control (Zmud 1998). However, in recent years, there has been a shift in editorial perspective toward recognizing the multifaceted role of reviewers. For instance, Lee (1995), an editor at *Management Information Systems Quarterly* (*MISQ*), highlighted the importance of reviewers articulating manuscript strengths and expressing negative criticisms in a tactful manner, particularly when deciding to reject a submission. Lee contends that if criticism cannot be conveyed in a kind and constructive manner, it may not be worth expressing at all. Saunders (2005) advocates for reviewers to function as "diamond cutters," assisting authors in refining their papers, rather than adopting the role of a "gatekeeper" solely focused on identifying fatal flaws in a manuscript. Straub (2008) emphasizes the significance of reviewers paying attention to the intellectual novelty of papers, while Rai (2016) asserts that a qualified reviewer should adopt a constructive rather than a destructive approach in providing virtuous reviews.

Recognizing the potential impact of nice review comments on authors, even in the face of manuscript rejection (Rao and Bareham 2022), there is a prevailing belief within academia in the necessity and desire for constructive and courteous peer reviews (Chong 2021; Rao and Bareham 2022). Encouraging comments, including expressions of support, are viewed as instrumental in nurturing junior researchers, who may undergo a "grieving process" following the rejection of their initial manuscripts, particularly if lacking confidence and experience in scientific publishing (Kotsis and Chung 2014; Rao and Bareham 2022). Additionally, female scientists may experience greater setbacks in terms of publication and career development following unprofessional reviews compared to their male counterparts.

Understanding the role of praise in peer review is essential for maintaining research integrity. Research integrity involves not only conducting research honestly and rigorously but also communicating findings transparently and without bias (Qiu et al. 2022). Inadequate or biased reviews can undermine the quality, fairness, and integrity of the peer review process (Resnik and Elmore 2016). To deepen our understanding of praise in peer review, this study utilizes data from the *British Medical Journal (BMJ)* to explore praise comments and their associated factors. To the best of our knowledge, this

represents the first study specifically focusing on praise in peer review. Diverging from previous descriptive studies, this paper delves into the factors linked to praise and elucidates the relationships between praise, cultural distance, and gender.

Our contributions extend to two main strands of literature. Firstly, this study enriches the existing literature on praise in peer review by documenting the social meaning of review comments, particularly emphasizing the role of nice comments and praise. Secondly, we contribute to the literature on the factors influencing praise by demonstrating that authors are more likely to receive praise when culturally closer to reviewers, and female reviewers are inclined to offer more praise.

The subsequent sections of this paper are organized as follows: Section 2 provides a review of existing studies related to peer review comments, elucidating on potential factors associated with the expression of praise, including cultural distance and gender. Section 3 outlines the data sources and methodologies employed. Section 4 conducts empirical studies to assess relationships between factors in praise. Finally, Section 5 concludes with a discussion of tentative explanations, implications, and limitations.

2. Related works

2.1. *The analysis of peer review comments*

The increasing accessibility of data pertaining to peer review has led to a burgeoning interest in peer review research (Bolek et al. 2023). The scholarly focus has shifted toward investigating the textual characteristics of peer review and the factors influencing these attributes. Various studies have delved into discerning features inherent in peer review comments. Notably, the length of reviews has emerged as an important indicator in the investigation of peer review. Several scholarly inquiries posit that review length, serving as a proxy for the quantity of information conveyed by reviewers to editors and authors, may serve as an indicative measure of review quality (Casnici et al. 2017; Geldsetzer et al. 2023). Bianchi, Grimaldo, and Squazzoni (2019) contend that review length stands as a valuable metric for evaluating reviewers. However, several studies indicate that significant caution must be taken when interpreting reviews of reviews as indicators of review quality. Xu et al. (2019) and Zhang, Wang, et al. (2022) argue that the relationship between review length and quality is contentious and instead propose a more modest point: that review length may at least reflect the time and effort invested by reviewers in the process. Goldberg et al. (2023) design experiments to test whether artificially lengthening reviews can create false perceptions of higher quality by manipulating review length. Their findings reveal that review length does not accurately represent review quality.

Moreover, the literature explores various factors related to review characteristics (Bravo et al. 2019; Buljan et al. 2020), with attention directed toward economic institutions or the characteristics of reviewers (Garcia et al. 2022; Garcia, Rodriguez-Sánchez, and Fdez-Valdivia 2019). For instance, Zhang, Wang, et al. (2022) suggest that review length is reflective of reviewers' efforts and undertake an exploration of the factors associated with review length. Subsequent studies delve into the emotional aspects of review comments. Utilizing review comments sourced from the *British Medical Journal*, Zhang, Wang, et al. (2022) reveal the communicative function of the term "however" in scientific discourse. It was observed that reviewers often commence peer reviews with commendations before transitioning to critiques.

2.2. *Culture distance and praise*

A body of research has undertaken an examination of macro-level factors associated with peer review, revealing the presence of biases at both the country and institution levels. These biases encompass country bias (Fox, Meyer, and Aimé 2023; Rubin, Rubin, and Segal 2023; Smith et al. 2023) and racial inequality (Liu, Rahwan, and AlShebli 2023). In addition to formal factors, informal elements have been recognized as influential in the realm of science (Crescenzi, Nathan, and Rodríguez-Pose 2016). For instance, Khosrowjerdi and Bornmann (2021) have identified a correlation between national culture and research activities, specifically noting that the self-expression values of nations are linked to their scientific performance. In terms of culture, Zhang, Wang, et al. (2022) have demonstrated that reviewers influenced by Confucianism tend to provide shorter reviews due to the unique high-context communication prevalent in East Asian countries.

Culture is defined as the shared beliefs, values, and practices characteristic of ethnic, religious, and social groups (Guiso, Sapienza, and Zingales 2006). Cultural proximity refers to the degree of similarity between the interacting parties. Cultural distinctions among countries are commonly quantified through the assessment of cultural distance. Cultural distance, as defined by Hofstede, refers to the extent to which the cultural norms, values, and practices of one country differ from those of another (Guan and Yan 2016; Hofstede 2011; Khosrowjerdi and Bornmann 2021). It is a concept used to quantify the differences between cultures, which can significantly impact communication and decision-making among actors, leading to social inequality (Berger, Cohen, and Zelditch 1972; Brett and Okumura 1998; Giannetti and Yafeh 2012). Cultural distance is often calculated by comparing six dimensions across different countries, including power distance index, individualism, uncertainty avoidance index, masculinity, long-term orientation, and indulgence vs. restraint. Researchers have increasingly

focused on the inequality engendered by cultural factors within the peer review processes (Murray et al. 2018; Rubin, Rubin, and Segal 2023).

On one hand, the proclivity for individuals to accept ideas and thoughts similar to their own has been noted (Teplitskiy et al. 2018), as similarity can capture attention and bridge social distances (Munz, Jung, and Alter 2020). Racial reciprocity in peer review has been elucidated (Demarest, Freeman, and Sugimoto 2014), and Huntington (2011) posited that cultural commonality fosters cooperation, while cultural differences may lead to divisions and conflicts. Cultural distances between involved parties entail information and effort costs, whereby culturally distant receivers face high communication costs to accept divergent thoughts (Kleimeier and Chaudhry 2015). Agrawal, Kapur, and McHale (2008) found that social proximity, particularly shared ethnicity, can facilitate knowledge flows, and relationships like co-culture or co-ethnicity can enhance reciprocity possibilities by reducing trust establishment costs (Kalnins and Chung 2006; Saxenian 2002). Decision makers, for instance, tend to offer greater advantages (Giannetti and Yafeh 2012) and lower thresholds (Kleimeier and Chaudhry 2015) to culturally similar partners.

On the other hand, "particularistic factors" unrelated to the function of science are prevalent in scientific practices. Particularly when confronted with uncertainty, individuals tend to evaluate social characters that differentiate certain individuals from others, such as gender and country of origin, before forming their evaluations (Adair 1982; Cole 1983). Murray et al. (2018) illustrate that limited opportunities for country homophily may disadvantage researchers from countries with lower scientific output. However, if the distribution of similarity between authors and reviewers is uneven in science, reciprocity may exacerbate inequality (Merton 1968). Consequently, under conditions of similar geographical proximity, dyad countries with similar cultures are anticipated to receive more favorable comments than those with divergent cultures.

Based on these considerations, we hypothesize that cultural distance is associated with praise, leading to our first hypothesis:

H1: Reviewers who are culturally closer to authors are more likely to provide longer praise, as reflected in a greater word count.

2.3. *Gender and praise*

Despite concerns surrounding the underrepresentation of women in the field of science (Fox et al. 2016; Moss-Racusin et al. 2012; Witteman et al. 2019), it is imperative to acknowledge the substantial and impactful

contributions made by female scientists in scientific endeavors. Beyond the confines of scientific disciplines, women exhibit distinct attitudes and behaviors that transcend various domains, such as law (King and Greening 2007), politics (Skjeie 1991), medical treatment (Roter and Hall 2004; Tsugawa et al. 2017), and public service (Miller and Segal 2019). Existing research indicates that women tend to display greater communicativeness (Roter and Hall 2004) and lower levels of overconfidence (Wallis et al. 2022, 2023) compared to their male counterparts. Furthermore, women demonstrate a propensity for higher interpersonal sensitivity (Chan et al. 2023), higher ethical standards (Nguyen et al. 2008), enhanced inclusivity (Ni et al. 2021), and a greater respect for others' emotions and thoughts (Hall 1978). The supportive and encouraging nature of women also facilitates a more receptive attitude toward criticism (Bradley et al. 2012) and enhances the benefits derived from their suggestions (Edmondson 1999).

Gender role theory posits that societal expectations regarding the roles of men and women differ, with men being anticipated to pursue career success and women to assume caregiving roles (Frone and Rice 1987; Gutek, Nakamura, and Nieva 1981). Consequently, this societal framework contributes to observed differences in interpersonal behavior and professional conduct between genders.

In the context of the peer-review system, despite the underrepresentation of women (Fox et al. 2016; Helmer et al. 2017), empirical evidence suggests that female editors positively impact article quality (Bransch and Kvasnicka 2022) and may provide more favorable assessments (Wing et al. 2010). Building upon these observations, we posit that gender is potentially related to the praise received within the peer-review process, leading to the formulation of the second hypothesis.

H2: Female reviewers are more likely to provide longer praise, as reflected in a greater word count.

3. Data and method

Since September 2014, manuscripts submitted to *BMJ* have disclosed their publication history on bmj.com following acceptance, encompassing manuscripts, review comments, and authors' responses (Zhang et al. 2023; Zhang, Wang, et al. 2022). The *BMJ* employs a fully open peer review process for all articles, meaning that both authors and reviewers are aware of each other's identities. The dataset for this study comprises 3,917 PDFs derived from 690 articles published between 2015 and 2020 on bmj.com. In light of the

potential correlation between the number of review rounds and the depth of review scrutiny (Sikdar et al. 2017) and attention received (Zhang et al. 2023), we have restricted our analysis to include only 301 articles that underwent the initial review round.

During the data processing phase, pertinent information such as comments, reviewer names, job titles, and institutions were extracted from referee reports. The process of name disambiguation was facilitated by the reviewers' signed names. To ascertain the gender of the reviewers, the AMiner Gender Prediction API was employed (Zhang, Wang, et al. 2022; Zhang, Xu, et al. 2022). Additionally, following the methodology of Zhang, Wang, et al. (2022, 2023), we collected the number of citations of the corresponding author prior to the publication of the paper from Scopus.

Drawing from previous research, we opted for manual content analysis, deeming it more reliable than sentiment analysis algorithms for peer review scrutiny (Luo et al. 2021; Wolfram, Wang, and Abuzahra 2021). Specifically, praise comments were identified through manual categorization. For the determination of reviewer country, we attributed it to the location of the reviewer's institution after extracting relevant details. Additionally, articles originating from England, Scotland, Northern Ireland, and Wales were re-classified as originating from the United Kingdom (UK). After excluding comments with missing variables, we retained 952 valid individual reviews, along with the relevant variables.

To ensure consistency with the analysis, we excluded papers where the cultural distance between authors and reviewers could not be measured. Table 1 presents the distribution of comments per paper, ranging from 1 to 7, with the vast majority of accepted manuscripts receiving between two and four comments (accounting for 84%).

Table 2 presents the attributes of the dataset. The average and median number of reviewers per manuscript are 3.26 and 3.00, respectively. Male reviewers constitute 68% of the sample (644 reviewers), while female reviewers account for 32% (308 reviewers). Senior researchers make up 52% of the reviewers (494 reviewers), with non-senior reviewers comprising 48% (458 reviewers). The average and median

Table 1. The quantity distribution of comments per paper.

No. of comments	Frequency
1	7
2	90
3	96
4	64
5	24
6	11
7	5

Table 2. Attributions of dataset.

Variable	Mean	Median	Number
No. of Reviewers per paper	3.26	3	-
Ratio of female reviewers	0.32	-	308
Ratio of senior researcher reviewers	0.52	-	494
Length of praise	36.13	27	-
Length of comments	562.49	485	-
Ratio (Length of praise/Length of comments)	0.11	0.06	-

lengths of praise in the comments are 36.13 words and 27 words, respectively. The average and median lengths of comments are 562.49 words and 485 words, respectively. The average and median ratios, calculated by dividing the length of praise by the length of the comments, are 0.11 and 0.06, respectively.

3.1. *Dependent variable*

While the study of praise in review comments has been relatively unexplored, research on evaluative texts, such as recommendation letters, has shown that the length of these letters is a significant indicator (Dutt et al. 2016; Oslock et al. 2024). In previous study, authors (Zhang, Wang, et al. 2022) have demonstrated that positive words are mainly concentrated in the overall evaluation. This is because reviewers are more inclined to offer specific critiques of methods or results rather than praise. Consequently, this study focuses on the overall praise. We conducted the manual labeling and extraction of peer review comments in a manner consistent with established procedures. In accordance with the coding method employed by Zheng et al. (2023) and Dutt et al. (2016), comments classified as "praise" were those that conveyed a positive or excellent tone. These comments typically expressed the reviewer's gratitude, positivity, enthusiasm, interest, admiration, approval, or respect, and highlighted the manuscript's value, novelty, or significance. For instance, a review comment that appraised a manuscript as follows: "This is a well-written, unique review on an important topic to all physicians and policymakers" was categorized as a praise comment. To enhance the reliability of the coding results (Tang, Wang, and Hu 2023), two researchers independently scrutinized these peer review comments and categorized the texts into respective variables. The reviewer's identity was anonymized during the coding process. Subsequently, a third researcher cross-validated and finalized the coding through consensus. The quantification of words in each comment expressing praise was then computed as the comment's length (Zhang, Wang, et al. 2022; Zhang, Xu, et al. 2022).

3.2. Measures

3.2.1. Independent variables

3.2.1.1. Cultural distance. Cultural distance, as conceptualized in this study, is rooted in Hofstede's 6-D model, encompassing power distance, uncertainty avoidance, individualism, individualism versus collectivism, masculinity versus femininity, and long-term versus short-term orientation, as well as indulgence versus restraint (Hofstede and Hofstede 2010). Utilizing Hofstede's cultural indices (Hofstede 1980), Kogut and Singh (1988) subsequently proposed a classic index derived from the deviations along each cultural dimension between two countries. These deviations were corrected for differences in the variances of each dimension and then arithmetically averaged, constituting the basis for the construction of the cultural distance index applied in this study.

The formula for calculating the cultural distance index is as follows:

$$Culdist_{ij} = 1/6 \sum_{k=1}^{6} \left[\left(I_{ki} - I_{kj} \right)^2 V_k \right] \tag{1}$$

where I_{ki} stands for the index for the *k-th* cultural dimension and *i-th* country (reviewer), I_{kj} stands for the index for the *k-th* cultural dimension and *j-th* country (corresponding author), and $Culdist_{ij}$ denotes the cultural difference between the *i-th* country (reviewer) and the *j-th* country (corresponding author), V_k is the variance of the index of the *k-th* dimension. In the case of multiple corresponding authors, the one with the minimum cultural distance from the reviewer among them is selected for consideration in the calculation.

Gender. Gender is operationalized as a dichotomous variable in this study. Specifically, this binary variable *Female* is coded as "1" to indicate the reviewer being a woman and "0" otherwise.

3.2.2. Control variables

3.2.2.2 Geographical distance. In order to mitigate the potential impact of geographical distance (Agrawal, Kapur, and McHale 2008; Petruzzelli 2011), we incorporate a variable representing the geographical distance between the capital cities of the corresponding author's and the reviewer's countries. Similar to our approach with cultural distance, we used the same corresponding author in calculating geographical distance. This distance information was obtained from the dataset provided by "Research and Expertise on the World Economy" (http://www.cepii.fr/PDF_PUB/wp/2011/wp2011-25.pdf). The natural logarithm of this variable is employed to enhance comparability in the subsequent analysis.

3.3. *Author prestige*

Several empirical studies suggest that papers authored by prestigious individuals are more likely to receive favorable evaluations (Huber et al. 2022; Okike et al. 2016). We have therefore incorporated author prestige into our analysis. Specifically, in line with the approaches of So (1998) and Guba and Tsivinskaya (2023), we measure academic prestige using the number of citations accrued by the corresponding author prior to the paper's publication. In cases where multiple corresponding authors are present, we select the one with the highest number of citations, as our intention is to control for the prestige conferred by the most prominent author.

3.4. *Recommendation*

Buljan et al. (2020) have demonstrated that reviewer recommendations exert a substantial influence on the linguistic attributes of review comments. Consequently, we have incorporated a recommendation variable into our analysis. The variable, *Recommendation*, has been manually encoded, with a value of 1 assigned if the manuscript received a recommendation for acceptance from the reviewer, and 0 otherwise. For instance, statements such as "The paper should be considered for publication" or "It is certainly worthy of publication" are indicative of a positive recommendation and have been accordingly coded.

3.5. *Academic title*

In order to control potential influences stemming from academic status on review comments, in line with prior research findings (Sun et al. 2023; Zhang, Wang, et al. 2022), we have incorporated a variable representing the academic status of the reviewer (Zhang, Wang, et al. 2022). This binary variable, *Senior*, is encoded as 1 if the reviewer holds a senior researcher position and 0 otherwise.

3.6. *Same country*

This variable, *Samecu*, commonly referred to as a dummy variable, is binary in nature. Specifically, it takes the value of 1 if both the reviewer and the corresponding author are situated in the same country, and 0 otherwise.

Table 3. Variable descriptions.

Construct	Variable	Type	Description
Dependent	Praise	count	Length of praise among each review
Independent	Culdist	count	Culture distance between corresponding author and reviewer
	Female	dummy	1 if the reviewer was female; 0 otherwise
Control	lnGeodist	count	Ln(Geographical distance of capital city of corresponding author's and reviewer's countries)
	lnAucit	count	Ln(Number of prior citations by corresponding author in Scopus+1)
	Recommendation	dummy	1 if the paper was recommended to accept; 0 otherwise
	Senior	dummy	1 if the reviewer was senior researcher; 0 otherwise
	Samecu	dummy	1 if the reviewer and corresponding author was in same country; 0 otherwise
	Length	count	Length of each review
	Year	dummy	Publication year

Table 4. Summary of descriptive statistics.

Variable	N	Mean	SD	Min	Max
Praise	952	36.12	34.77	0.00	341.00
Culdist	952	0.93	1.12	0.00	6.23
Female	952	0.32	0.47	0.00	1.00
lnGeodist	952	7.64	1.43	4.34	9.86
lnAucit	952	6.66	2.70	0.00	11.99
Recommendation	952	0.14	0.34	0.00	1.00
Senior	952	0.52	0.50	0.00	1.00
Samecu	952	0.25	0.43	0.00	1.00
Length	952	562.49	390.78	1.00	2535.00
Year	952	2017.37	1.50	2015.00	2020.00

Table 5. Spearman correlation matrix.

	Variable	1	2	3	4	5	6	7	8	9	10
1	Praise	1.00									
2	Culdist	−0.08	1.00								
3	Female	0.04	−0.01	1.00							
4	lnGeodist	−0.04	0.33	0.02	1.00						
5	lnAucit	0.03	−0.07	−0.00	−0.04	1.00					
6	Recommendation	0.20	0.00	−0.03	−0.04	−0.01	1.00				
7	Senior	−0.00	−0.01	−0.16	0.03	0.04	−0.07	1.00			
8	Samecu	0.03	−0.47	0.02	−0.59	0.05	−0.02	0.01	1.00		
9	Length	0.06	−0.05	−0.03	−0.01	−0.02	−0.05	−0.03	0.00	1.00	
10	Year	0.00	0.13	0.09	0.05	0.05	−0.06	−0.05	−0.00	−0.00	1.00

3.7. *Length*

This variable, *Length*, signifies the length of each review.

Table 3 delineates the definitions and sources of the variables. The descriptive statistics and pairwise correlation matrix for these variables are presented in Tables 4 and 5, respectively.

3.8. *Method*

Given that *Praise* is a discrete variable composed of non-negative integers, we employed Negative Binomial Regression to examine the association between *Praise* and various factors. Following the exclusion of observations with missing data for relevant variables, a total of 952 observations were retained for inclusion in the regression analysis. The statistical analysis was conducted using Stata version 15.1.

Our model is specified in Eq. (2) below, where the dependent variable is denoted as *Praise_i*. The coefficient β_1 represents the relationship between cultural distance and praise, while β_2 captures the relationship between the reviewer's gender and praise.

$$Praise_i = \alpha + \beta_1 Culdist_i + \beta_2 Female_i + Other\ control\ variables_i + \varepsilon i \quad (2)$$

ε_i is an error term.

4. Results

As depicted in Table 4, the mean length of praise for each review comment is 36.12 words. The average cultural distance between two countries is 0.93, while the average natural logarithm of geographical distance (ln(Geographical distance)) is 7.64. The female reviewers constitute 32% (308) of the sample. Approximately 14% (129) of the review comments recommend the acceptance of manuscripts. Additionally, 52% (494) of the reviewers hold senior researcher positions, and 25% (235) share the same country of origin as the corresponding author. The average length of the review comments is 562.49 words. An examination of Table 5 reveals minimal correlation among the variables under consideration.

4.1. *Benchmark models*

Table 6 presents the outcomes of the regression model articulated in Eq. (2). In Columns 1 and 2, only the independent variables *Culdist* and *Female* are respectively included. Column 1 demonstrates that *Culdist* is significantly associated with *Praise*, indicating that authors are more likely to receive praise when they are culturally closer to reviewers. While the coefficients are positive, Column 2 indicates that there is insignificant relationship between *Female* and *Praise*.

Control variables are introduced in Columns 3 and 4, where the magnitude and significance of coefficients exhibit no substantial changes. All variables are incorporated in Column 5. The coefficient of *Culdist* is negative and significant. These findings provide preliminary support for our H1.

Table 6. Regression results.

	(1)	(2)	(3)	(4)	(5)
	Culdist	Female	Culdist	Female	All
Culdist	−0.078***		−0.078**		−0.074**
	(0.027)		(0.031)		(0.031)
Female		0.076		0.127**	0.116*
		(0.064)		(0.065)	(0.064)
lnGeodist			−0.003	−0.006	−0.005
			(0.026)	(0.027)	(0.026)
lnAuCit			0.006	0.010	0.007
			(0.011)	(0.011)	(0.011)
Recommendation			0.480***	0.495***	0.494***
			(0.079)	(0.080)	(0.080)
Senior			0.018	0.038	0.035
			(0.061)	(0.061)	(0.061)
Samecu			−0.016	0.066	−0.019
			(0.097)	(0.089)	(0.096)
Length			0.000**	0.000**	0.000**
			(0.000)	(0.000)	(0.000)
Constant	3.629***	3.543***	3.379***	3.239***	3.339***
	(0.091)	(0.090)	(0.259)	(0.255)	(0.258)
Observations	952	952	952	952	952
ll	−4356.47	−4358.49	−4344.36	−4345.14	−4343.38
aic	8728.95	8732.99	8716.73	8718.28	8716.76
bic	8767.82	8771.86	8784.75	8786.30	8789.64
year	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. In the regression analysis that excluded papers with multiple corresponding authors, the results remained unaffected.

However, the coefficient of *Female* is significant in column 4 but is insignificant in column 5.

Regarding the control variables, in contrast to cultural distance, there exist no discernible relationships between *lnGeodist* and *Praise*. Likewise, the shared countries between reviewers and corresponding authors, as well as the status of being senior researchers or corresponding author's prestige, does not yield any significant effects on the length of praise compared to other reviewers. Comments recommending manuscripts tend to result in longer expressions of praise.

Given the significant impact of reviewer recommendations on the linguistic characteristics of review comments, we further divided the observations into two categories: *Recommended* and *Non-Recommended*, based on the *Recommendation* variable and conducted new regressions (see Table 7). Columns 1 and 3 represent the *Recommended* category, while Columns 2 and 4 represent the *Non-Recommended* category. We incorporate the variable *Female* in Columns 1 and 2, and include all variables in Columns 3 and 4.

The results reveal divergent relationships between gender and *Praise*. Specifically, for papers not recommended for acceptance, female reviewers exhibit a tendency to provide longer expressions of praise. However, for

Table 7. Regression results for subgroups.

	(1)	(2)	(3)	(4)	(5)	(6)
	Recommended	Non-Recommended	Recommended	Non-Recommended	Re-Au	Re-Au
Culdist			−0.074	−0.075**		−0.073**
			(0.081)	(0.033)		(0.032)
Female	−0.333**	0.171**	−0.260*	0.168**		
	(0.147)	(0.070)	(0.139)	(0.070)		
Re(Fe)-Au(M)					−0.066	−0.048
					(0.081)	(0.079)
Re(M)-Au(F)					0.010	0.059
					(0.081)	(0.082)
Re(F)-Au(F)					0.111	0.151
					(0.095)	(0.095)
lnGeodist			−0.123**	0.024		−0.004
			(0.055)	(0.029)		(0.026)
lnAuCit			0.044	0.001		0.008
			(0.028)	(0.012)		(0.012)
Senior			0.088	0.032		0.033
			(0.142)	(0.067)		(0.061)
Samecu			−0.431**	0.049		−0.019
			(0.178)	(0.107)		(0.097)
Length			0.000	0.000*		0.000**
			(0.000)	(0.000)		(0.000)
Recommendation						0.495***
						(0.080)
Constant	4.049***	3.355***	4.697***	3.112***	3.571***	3.352***
	(0.135)	(0.108)	(0.542)	(0.288)	(0.094)	(0.264)
Observations	129	823	129	823	948	948
loglikelihood	−631.62	−3691.93	−624.75	−3688.59	−4339.04	−4324.15
aic	1279.23	7399.86	1277.50	7405.18	8698.08	8682.31
bic	1302.11	7437.56	1317.54	7471.16	8746.63	8764.83
Year	Yes	Yes	Yes	Yes	Yes	Yes

papers recommended for acceptance, female reviewers appear to offer shorter praise. This discrepancy may contribute to the non-statistically significant coefficient of the female variable in Table 6. Consequently, our findings lend partial support to H2.

Additionally, we introduced reviewer-author pairs into the analysis to explore whether female authors receive more favorable reviews from female reviewers, specifically in columns 5–6. We created a dummy variable to capture the different combinations, including Re(M)-Au (M), Re(F)-Au(M), Re(M)-Au(F), and Re(F)-Au(F). For instance, the Re(M)-Au(M) group, which serves as the baseline in the regression, represents instances where both the corresponding author and reviewer are male. We excluded papers with multiple corresponding authors of differing genders. Column 5 includes only the reviewer-author pairs variable, while Column 6 incorporates all variables. Our analysis did

Table 8. Regression results for robustness tests.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Non UK	Culdist'	Length	Poisson	IV	PSM	Non_Im
Culdist	−0.068**		−0.041*	−0.076**	−0.124**	−0.088**	−0.074**
	(0.032)		(0.022)	(0.031)	(0.052)	(0.037)	(0.035)
Culdist'		−0.058**					
		(0.028)					
Female	0.142**	0.117*	−0.049	0.099	0.095	0.152**	0.077
	(0.071)	(0.064)	(0.048)	(0.062)	(0.075)	(0.071)	(0.078)
lnGeodist	−0.009	−0.008	−0.005	−0.018	−0.016	−0.026	−0.032
	(0.029)	(0.026)	(0.019)	(0.028)	(0.029)	(0.029)	(0.030)
lnAuCit	0.009	0.008	−0.007	0.010	0.011	0.003	0.010
	(0.012)	(0.011)	(0.008)	(0.012)	(0.013)	(0.012)	(0.014)
Recommendation	0.420***	0.497***	−0.119	0.483***	0.509***	0.322***	0.535***
	(0.083)	(0.080)	(0.075)	(0.081)	(0.093)	(0.096)	(0.093)
Senior	0.070	0.034	−0.054	0.050	0.016	−0.010	−0.006
	(0.067)	(0.061)	(0.045)	(0.062)	(0.071)	(0.076)	(0.074)
Samecu	−0.019	−0.073	−0.062	−0.057		−0.120	
	(0.104)	(0.114)	(0.068)	(0.097)		(0.113)	
Length	0.000*	0.000**		0.000**	0.000*	0.000*	0.000*
	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)
Constant	3.372***	3.409***	6.513***	3.451***	7.974	3.533***	3.615***
	(0.286)	(0.270)	(0.186)	(0.253)	(50.284)	(0.298)	(0.301)
Observations	725	952	952	952	707	616	630
loglikelihood	−3323.02	−4343.84	−6852.48	−15832.63	-	−2809.26	−2860.96
aic	6676.05	8717.68	13732.96	31693.26	-	5648.53	5749.92
bic	6744.84	8790.56	13800.98	31761.28	-	5714.88	5812.16
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes

not find evidence that female authors receive more positive reviews from female reviewers.

We further provide a series of results of robustness tests to enhance the reliability of conclusions (see Tables 8 and 9).

4.2. Robustness tests

Acknowledging the potential presence of home bias, we exclude reviewers working in the United Kingdom in Column 1 (see Table 8). In Column 2, an alternative method to measure cultural distance is employed, following the approach proposed by Vermeulen and Barkema (2001). This method utilizes a Euclidean index rather than the arithmetically averaged index (Kogut and Singh 1988) based on Hofstede's (1980) cultural dimensions, as depicted in Eq. (3). In contrast to Kogut and Singh (1988), this method posits that the various cultural dimensions determining scores of cultural distance are not equally important.

Table 9. Regression results for robustness tests (non-recommended).

	(8) Non UK	(9) Culdist'	(10) Length	(11) Poisson	(12) Non_Im
Culdist	−0.066*		−0.031	−0.081**	−0.075**
	(0.034)		(0.022)	(0.034)	(0.037)
Culdist'		−0.062**			
		(0.031)			
Female	0.206***	0.162**	−0.053	0.175**	0.178**
	(0.077)	(0.070)	(0.050)	(0.069)	(0.074)
lnGeodist	0.013	0.022	−0.011	0.027	0.004
	(0.031)	(0.029)	(0.021)	(0.030)	(0.030)
lnAucit	0.010	0.002	−0.011	−0.000	−0.003
	(0.014)	(0.012)	(0.009)	(0.012)	(0.013)
Senior	0.099		−0.055	0.038	−0.012
	(0.074)		(0.046)	(0.066)	(0.072)
Samecu	0.048	−0.011	−0.064	0.061	−0.023
	(0.115)	(0.126)	(0.071)	(0.111)	(0.118)
Length	0.000	0.000*		0.000**	0.000*
	(0.000)	(0.000)		(0.000)	(0.000)
Constant	3.061***	3.205***	6.577***	3.096***	3.331***
	(0.321)	(0.299)	(0.195)	(0.304)	(0.302)
Observations	624	823	823	823	719
loglikelihood	−2818.01	−3688.94	−5915.42	−13463.03	−3208.46
aic	5664.02	7403.87	11856.84	26952.07	6444.92
bic	5726.13	7465.14	11918.11	27013.34	6509.01
Year	Yes	Yes	Yes	Yes	Yes

$$Culdist'_{ij} = \sqrt{\sum_{k=1}^{6}\left[\left(I_{ki} - I_{kj}\right)^2/V_k\right]} \tag{3}$$

Here, I_{ki} denotes the index for the *k-th* cultural dimension and the *i-th* country (reviewer), I_{kj} stands for the index for the *k-th* cultural dimension and the *j-th* country (corresponding author), $Culdist'_{ij}$ represents the cultural difference between the *i-th* country, where the reviewer works, and the *j-th* country, where the corresponding author works, and V_k denotes the variance of the index of the *k-th* dimension. The coefficients of *Culdist* and *Culdist'* are positively significant in both Column 1 and Column 2.

We introduce the length of each review as a "placebo" in Column 3, altering the regression model to use Poisson regression in Column 4. The coefficients of *Culdist* are found to be statistically insignificant in Column 3 but exhibit statistical significance in Column 4. To address potential endogeneity concerns with cultural distance, we employ the difference in religious beliefs as an instrumental variable (IV). The difference in religious beliefs is correlated with cultural distance but lacks an independent effect on *Praise*. We categorize countries into six groups, namely Catholic, Protestantism, Confucianism, Muslim, Judaism, and other (Gill & Lundsgaarde, 2004). This IV is assigned a value of 1 if the reviewer and corresponding author share the same religious beliefs country, and 0 otherwise. We retain observations where

reviewers and corresponding authors are in different countries and utilize IV/GMM Poisson regression in Column 5. The results in Column 5 demonstrate that the coefficient of cultural distance is negative and significant.

Given that the gender variable is dichotomous, we employ Propensity Score Matching (PSM) and conduct a new regression in Column 6. Specifically, control variables serve as the basis for matching. In addition, reviewers may reside and work in a country different from their country of origin. For instance, a reviewer might be employed in the United Kingdom but originally from China, potentially introducing bias in the measurement of "cultural distance." Following the methodology outlined by Zhang, Wang, et al. (2022), we utilized a surname dictionary to match the remaining surnames with countries of origin. Subsequently, we excluded reviewers working in countries other than their countries of origin and performed a new regression analysis, as presented in Column 7.

We replicate the procedures outlined in (1), (2), (3), (4), and (7) at (8)–(12) for Non-Recommended reviews (see Table 9). The results in Tables 8 and 9 reveal that the magnitude and significance of variables align consistently with the previous regression results.

5. Discussion and conclusion

5.1. *Discussion*

To the best of our knowledge, and in contrast with existing studies on review comments (Zhang, Wang, et al. 2022; Zhang, Xu, et al. 2022), our study undertakes a quantitative investigation into the praise in peer review. This analysis contributes to the growing body of literature by introducing novel factors associated with praise in peer review. Utilizing data from *BMJ*, our findings reveal that macro-level homophilic interactions and individual characteristics are correlated with praise. Specifically, our results show that authors who are culturally closer to reviewers are more likely to receive praise, and female reviewers exhibit a greater propensity to offer praise. Aligning with previous research (Black et al. 1998; Card and DellaVigna 2020; Sun et al. 2023), we observe no discernible relationship between reviewers' academic status and praise in peer review. To foster a more constructive peer review process, it is essential to implement structured review protocols and encourage constructive comments. Unlike a purely critical review, which might only point out flaws or reasons for rejection, constructive peer review aims to offer specific actionable suggestions that the author can use to enhance the quality of their manuscript. Additionally, an inclusive research culture refers to an environment where researchers from all backgrounds feel valued, respected, and supported, and where diverse

perspectives are integrated into the research process. Promoting a more inclusive research culture necessitates increasing the diversity of reviewers, particularly from underrepresented demographics.

The potential for cultural similarity and gender to influence reviewer demographics raises concerns about the uneven distribution of these factors among reviewers. Notably, reviewers from Anglosphere nations, such as the United Kingdom, United States, Canada, Australia, and New Zealand, constitute a significant proportion of the reviewer pool (Fox, Meyer, and Aimé 2023; Rubin, Rubin, and Segal 2023; Smith et al. 2023). Evidence also indicates the underrepresentation of female reviewers (Fox et al. 2019; Helmer et al. 2017; Pinho-Gomes et al. 2021).

The composition of the reviewer group suggests that certain authors are more likely to face disadvantages in peer review, especially those from countries with greater cultural diversity. Authors from specific regions, such as Asian East/Southeast nations, may find themselves reviewed by individuals from culturally different countries, potentially posing challenges, particularly for early career researchers.

To mitigate potential inequalities, it may be effective to implement structured peer review processes, provide clear guidelines for reviewers, and establish systematic methods for assessing the quality of review reports (Malički and Mehmani 2024). Providing reviewers with appropriate guidance may enhance the fairness and integrity of the peer review process. Efforts to enhance peer review should continue to focus on promoting diversity in terms of culture and gender among reviewers. Journal editors are urged to transcend convenience and expand their networks to include a more diverse pool of reviewers. Reviewers, as unsung heroes, bear a substantial burden of review work without compensation. Despite the inherent challenges and competitiveness in scientific research, scientists should strive to conduct science in a more compassionate manner. In addition, it is essential to recognize that positive comments, including praise, are not neutral and may potentially bias editorial decisions. Negative comments, however, often wield greater influence on individuals than positive remarks (Liebrecht, Hustinx, and van Mulken 2019).

As to method, we believe it is crucial to develop more refined methods for identifying the tone of review comments. In future analyses, we plan to explore the use of generative artificial intelligence, such as ChatGPT, to design an improved identification approach.

Finally, the interactions between reviewers and authors could be incorporated into the peer review model to better understand the process (Garcia et al. 2022; Garcia, Rodriguez-Sánchez, and Fdez-Valdivia 2020a, 2020b). Exploring the patterns of knowledge communication within peer review can enhance our understanding of its mechanisms. In practical terms, it is important to encourage journals to promote open peer review under specific

conditions and to offer greater recognition and affirmation to reviewers for their contributions.

6. Conclusion

This study represents the attempt to systematically investigate factors associated with praise in peer review. The findings indicate that authors are more likely to receive praise when they share closer cultural ties with reviewers. Additionally, female reviewers tend to provide a higher degree of praise.

In aggregate, these outcomes contribute valuable insights for fostering a more constructive and inclusive research culture, promoting positive interactions within peer review processes. The results may also highlight underlying inequalities in science, potentially extending across various scientific disciplines. Peer review extends beyond the realm of publications and encompasses contexts such as grant applications and job evaluations. Given that peer review serves not only as a critical tool for quality control but also as a platform for scholarly communication, uncovering potential biases and addressing inequalities across the wider peer review system is crucial for the scientific community.

Previous studies have demonstrated that removing anonymity can enhance the tone and constructiveness of reviews without compromising scientific or editorial value or diminishing the quality of peer review. Promoting greater transparency is helpful in ensuring the integrity of the research we publish. Therefore, this study provides useful guidance for editors and journals in supporting research integrity and creating equitable environments.

7. Limitation

It is important to acknowledge several limitations inherent in our study. Firstly, this study only focused on accepted papers, thus the findings are only applicable to such papers. Secondly, the data utilized for this investigation originates from *BMJ*, a journal renowned for its high quality and a reviewer workforce comprising numerous exceptional and conscientious individuals. Therefore, the generalizability of our conclusions may warrant scrutiny when applied to other datasets in the broader context of peer review. Future studies should consider peer review data from journals encompassing various subject areas and different impact factor quartiles. Thirdly, regarding the peer review model, although current evidence suggests that open review practices do not significantly affect the objectivity of the peer review process, reviewers tend to write more positive reports if they know their identities will be disclosed (Bravo et al. 2019). This introduces a potential bias into our study. Although there is no evidence to suggest that editors can edit reviewers' comments, potential bias could be introduced into the analysis if such editing were to occur. Finally, regarding

the method, we can only identify the recommendation for acceptance based on the reviewers' comments. However, we do not have access to the specific recommendation type each reviewer selected in the system when submitting their review report. The primary focus in this study is on the overall evaluation of praise. Future studies could also consider categorizing praise according to each specific aspect of a reviewer's comments.

Acknowledgments

Disclosure statement

Funding

References

Adair, R. K. 1982. "A Physics Editor Comments on Peters and Ceci's Peer-Review Study." *Behavioral and Brain Sciences* 5 (2): 196–196. https://doi.org/10.1017/S0140525X00011195.

Agrawal, A., D. Kapur, and J. McHale. 2008. "How Do Spatial and Social Proximity Influence Knowledge Flows? Evidence from Patent Data." *Journal of Urban Economics* 64 (2): 258–269. https://doi.org/10.1016/j.jue.2008.01.003.

Baglini, R., and C. Parsons. 2020. "If You can't Be Kind in Peer Review, Be Neutral." *Nature.* https://doi.org/10.1038/d41586-020-03394-y.

Bakanic, V., C. McPhail, and R. J. Simon. 1987. "The Manuscript Review and Decision-Making Process." *American Sociological Review* 52 (5): 631–642. https://doi.org/10.2307/2095599.

Berger, J., B. P. Cohen, and M. Zelditch. 1972. "Status Characteristics and Social Interaction." *American Sociological Review* 37 (3): 241–255. https://doi.org/10.2307/2093465.

Bharti, P. K., M. Navlakha, M. Agarwal, and A. Ekbal. 2023. "PolitePEER: Does Peer Review Hurt? A Dataset to Gauge Politeness Intensity in the Peer Reviews." *Language Resources and Evaluation.* https://doi.org/10.1007/s10579-023-09662-3.

Bianchi, F., F. Grimaldo, & F. Squazzoni. 2019. "The F-3-index. Valuing Reviewers for Scholarly Journals." *Journal of Informetrics* 13 (1): 78–86. https://doi.org/10.1016/j.joi.2018.11.007.

Black, N., S. van Rooyen, F. Godlee, R. Smith, and S. Evans. 1998. "What Makes a Good Reviewer and a Good Review for a General Medical Journal?" *Jama-Journal of the American Medical Association* 280 (3): 231–233. https://doi.org/10.1001/jama.280.3.231.

Bolek, M., C. Bolek, J. Shopovski, and D. Marolov. 2023. "The Consistency of Peer-Reviewers: Assessment of Separate Parts of the Manuscripts Vs Final Recommendations." *Accountability in Research* 30 (7): 493–515. https://doi.org/10.1080/08989621.2022.2030719.

Bradley, B. H., B. E. Postlethwaite, A. C. Klotz, M. R. Hamdani, and K. G. Brown. 2012. "Reaping the Benefits of Task Conflict in Teams: The Critical Role of Team Psychological Safety Climate." *Journal of Applied Psychology* 97 (1): 151–158. https://doi.org/10.1037/a0024200.

Bransch, F., and M. Kvasnicka. 2022. "Male Gatekeepers: Gender Bias in the Publishing Process?" *Journal of Economic Behavior and Organization* 202:714–732. https://doi.org/10.1016/j.jebo.2022.07.031.

Bravo, G., F. Grimaldo, E. López-Iñesta, B. Mehmani, and F. Squazzoni. 2019. "The Effect of Publishing Peer Review Reports on Referee Behavior in Five Scholarly Journals." *Nature Communications* 10 (1): 322. https://doi.org/10.1038/s41467-018-08250-2.

Brett, J. M., and T. Okumura. 1998. "Inter- and Intracultural Negotiation: U.S. and Japanese Negotiators." *Academy of Management Journal* 41 (5): 495–510. https://doi.org/10.2307/256938.

Buljan, I., D. Garcia-Costa, F. Grimaldo, F. Squazzoni, and A. Marušić. 2020. "Large-Scale Language Analysis of Peer Review Reports." *eLife* 9:e53249. https://doi.org/10.7554/eLife.53249.

Card, D., and S. DellaVigna. 2020. "What Do Editors Maximize? Evidence from Four Economics Journals." *The Review of Economics and Statistics* 102 (1): 195–217. https://doi.org/10.1162/rest_a_00839.

Casnici, N., F. Grimaldo, N. Gilbert, and F. Squazzoni. 2017. "Attitudes of Referees in a Multidisciplinary Journal: An Empirical Analysis." *Journal of the Association for Information Science and Technology* 68 (7): 1763–1771. https://doi.org/10.1002/asi.23665.

Chan, T. H., H. Liu, S. Keck, and W. Tang. 2023. "When Do Teams Generate Valuable Inventions? The Moderating Role of Invention Integrality on the Effects of Expertise Similarity, Network Cohesion, and Gender Diversity." *Production & Operations Management* 32 (6): 1760–1777. https://doi.org/10.1111/poms.13939.

Chong, S. W. 2021. "Improving Peer-Review by Developing reviewers' Feedback Literacy." *Learned Publishing* 34 (3): 461–467. https://doi.org/10.1002/leap.1378.

Clements, J. C. 2020. "Don't Be a Prig in Peer Review." *Nature* 585 (7825): 472–472. https://doi.org/10.1038/d41586-020-02512-0.

Cole, S. 1983. "The Hierarchy of the Sciences?" *The American Journal of Sociology* 89 (1): 111–139. https://doi.org/10.1086/227835.

Crescenzi, R., M. Nathan, and A. Rodríguez-Pose. 2016. "Do Inventors Talk to Strangers? On Proximity and Collaborative Knowledge Creation." *Research Policy* 45 (1): 177–194. https://doi.org/10.1016/j.respol.2015.07.003.

Crijns, T. J., J. S. E. Ottenhoff, and D. Ring. 2021. "The Effect of Peer Review on the Improvement of Rejected Manuscripts." *Accountability in Research-Policies and Quality Assurance* 28 (8): 517–527. https://doi.org/10.1080/08989621.2020.1869547.

Demarest, B., G. Freeman, and C. R. Sugimoto. 2014. "The Reviewer in the Mirror: Examining Gendered and Ethnicized Notions of Reciprocity in Peer Review." *Scientometrics* 101 (1): 717–735. https://doi.org/10.1007/s11192-014-1354-z.

Derrick, G. 2020. "How COVID-19 Lockdowns Could Lead to a Kinder Research Culture." *Nature* 581 (7806): 107–108. https://doi.org/10.1038/d41586-020-01144-8.

Derrick, G. 2021. "A Year After Lockdowns Began, Has Research Got Any Kinder?" *Nature*. https://doi.org/10.1038/d41586-021-01579-7.

Dutt, K., D. L. Pfaff, A. F. Bernstein, J. S. Dillard, and C. J. Block. 2016. "Gender Differences in Recommendation Letters for Postdoctoral Fellowships in Geoscience." *Nature Geoscience* 9 (11), 805–808. https://doi.org/10.1038/ngeo2819.

Edmondson, A. 1999. "Psychological Safety and Learning Behavior in Work Teams." *Administrative Science Quarterly* 44 (2): 350–383. https://doi.org/10.2307/2666999.

Fox, C. W., C. S. Burns, J. A. Meyer, and K. Thompson. 2016. "Editor and Reviewer Gender Influence the Peer Review Process but Not Peer Review Outcomes at an Ecology Journal." *Functional Ecology* 30 (1): 140–153. https://doi.org/10.1111/1365-2435.12529.

Fox, C. W., M. A. Duffy, D. J. Fairbairn, and J. A. Meyer. 2019. "Gender Diversity of Editorial Boards and Gender Differences in the Peer Review Process at Six Journals of Ecology and Evolution." *Ecology and Evolution* 9 (24): 13636–13649. https://doi.org/10.1002/ece3.5794.

Fox, C. W., J. Meyer, and E. Aimé. 2023. "Double-Blind Peer Review Affects Reviewer Ratings and Editor Decisions at an Ecology Journal." *Functional Ecology* 37 (5): 1144–1157. https://doi.org/10.1111/1365-2435.14259.

Frone, M., and R. Rice. 1987. "Work-Family Conflict: The Effect of Job and Family Involvement." *Journal of Organizational Behavior* 8 (1): 45–53. https://doi.org/10.1002/job.4030080106.

Garcia, J. A., J. Chamorro-Padial, R. Rodriguez-Sanchez, and J. Fdez-Valdivia. 2022. "What is the Sensitivity and Specificity of the Peer Review Process?" *Accountability in Research-Policies and Quality Assurance*: 1–22. https://doi.org/10.1080/08989621.2022.2122817.

Garcia, J. A., R. Rodriguez-Sánchez, and J. Fdez-Valdivia. 2019. "The Optimal Amount of Information to Provide in an Academic Manuscript." *Scientometrics* 121 (3): 1685–1705. https://doi.org/10.1007/s11192-019-03270-1.

Garcia, J. A., R. Rodriguez-Sánchez, and J. Fdez-Valdivia. 2020a. "The Author–Reviewer Game." *Scientometrics* 124 (3): 2409–2431. https://doi.org/10.1007/s11192-020-03559-6.

Garcia, J. A., R. Rodriguez-Sánchez, and J. Fdez-Valdivia. 2020b. "Confirmatory Bias in Peer Review." *Scientometrics* 123 (1): 517–533. https://doi.org/10.1007/s11192-020-03357-0.

Geldsetzer, P., M. Heemann, P. Tikka, G. Wang, M. M. Cusick, A. Lenjani, and N. Krishnan. 2023. "Prevalence of Short Peer Reviews in 3 Leading General Medical Journals." *JAMA Network Open* 6 (12): e2347607. https://doi.org/10.1001/jamanetworkopen.2023.47607.

Giannetti, M., and Y. Yafeh. 2012. "Do Cultural Differences Between Contracting Parties Matter? Evidence from Syndicated Bank Loans." *Management Science* 58 (2): 365–383. https://doi.org/10.1287/mnsc.1110.1378.

Gill, A., and E. Lundsgaarde. 2004. "State Welfare Spending and Religiosity: A Cross-National Analysis." *Rationality and Society* 16 (4): 399–436. https://doi.org/10.1177/1043463104046694.

Goldberg, A., I. Stelmakh, K. Cho, A. Oh, A. Agarwal, D. Belgrave, and N. B. Shah. 2023. "Peer Reviews of Peer Reviews: A Randomized Controlled Trial and Other Experiments." (arXiv:2311.09497). arXiv. https://doi.org/10.48550/arXiv.2311.09497.

Guan, J., and Y. Yan. 2016. "Technological Proximity and Recombinative Innovation in the Alternative Energy Field." *Research Policy* 45 (7): 1460–1473. https://doi.org/10.1016/j.respol.2016.05.002.

Guba, K., and A. Tsivinskaya. 2023. "Expert Judgments versus Publication-Based Metrics: Do the Two Methods Produce Identical Results in Measuring Academic Reputation?" *Journal of Documentation* 79 (1): 127–143. https://doi.org/10.1108/JD-02-2022-0039.

Guiso, L., P. Sapienza, and L. Zingales. 2006. "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives* 20 (2): 23–48. https://doi.org/10.1257/jep.20.2.23.

Gutek, B. A., C. Y. Nakamura, and V. F. Nieva. 1981. "The Interdependence of Work and Family Roles." *Journal of Organizational Behavior* 2 (1): 1–16. https://doi.org/10.1002/job.4030020102.

Hall, J. A. 1978. "Gender Effects in Decoding Nonverbal Cues." *Psychological Bulletin* 85 (4): 845–857. https://doi.org/10.1037/0033-2909.85.4.845.

Helmer, M., M. Schottdorf, A. Neef, and D. Battaglia. 2017. "Gender Bias in Scholarly Peer Review." *eLife* 6:e21718. https://doi.org/10.7554/eLife.21718.

Hofstede, G. 1980. *Culture's Consequences: International Differences in Work-Related Values*. Abridged ed. Beverly Hills, CA: SAGE Publications, Inc.

Hofstede, G. 2011. "Dimensionalizing Cultures: The Hofstede Model in Context." *Online Readings in Psychology and Culture* 2 (1). https://doi.org/10.9707/2307-0919.1014.

Hofstede, G., and G. J. Hofstede. 2010. *Cultures and Organizations, Software of the Mind*. New York: McGraw Hill Professional.

Huber, J., S. Inoua, R. Kerschbamer, C. König-Kersting, S. Palan, and V. L. Smith. 2022. "Nobel and Novice: Author Prominence Affects Peer Review." *Proceedings of the National Academy of Sciences* 119 (41): e2205779119. https://doi.org/10.1073/pnas.2205779119.

Huntington, S. P. 2011. *The Clash of Civilizations and the Remaking of World Order*. New York: Simon & Schuster.

Kalnins, A., and W. Chung. 2006. "Social Capital, Geography, and Survival: Gujarati Immigrant Entrepreneurs in the U.S. Lodging Industry." *Management Science* 52 (2): 233–247. https://doi.org/10.1287/mnsc.1050.0481.

Khosrowjerdi, M., and L. Bornmann. 2021. "Is Culture Related to Strong Science? An Empirical Investigation." *Journal of Informetrics* 15 (4): 101160. https://doi.org/10.1016/j.joi.2021.101160.

King, K. L., and M. Greening. 2007. "Gender Justice or Just Gender? The Role of Gender in Sexual Assault Decisions at the International Criminal Tribunal for the Former Yugoslavia*." *Social Science Quarterly* 88 (5): 1049–1071. https://doi.org/10.1111/j.1540-6237.2007.00491.x.

Kleimeier, S., and S. M. Chaudhry. 2015. "Cultural Differences and the Structure of Loan Syndicates." *Finance Research Letters* 15:115–124. https://doi.org/10.1016/j.frl.2015.09.001.

Kogut, B., and H. Singh. 1988. "The Effect of National Culture on the Choice of Entry Mode." *Journal of International Business Studies* 19 (3): 411–432. https://doi.org/10.1057/palgrave.jibs.8490394.

Kotsis, S. V., and K. C. Chung. 2014. "Manuscript Rejection: How to Submit a Revision and Tips on Being a Good Peer Reviewer." *Plastic and Reconstructive Surgery* 133 (4): 958. https://doi.org/10.1097/PRS.0000000000000002.

Lee, A. S. 1995. "Reviewing a Manuscript for Publication." *Journal of Operations Management* 13 (1): 87–92. https://doi.org/10.1016/0272-6963(95)94762-w.

Liebrecht, C., L. Hustinx, and M. van Mulken. 2019. "The Relative Power of Negativity: The Influence of Language Intensity on Perceived Strength." *Journal of Language & Social Psychology* 38 (2): 170–193. https://doi.org/10.1177/0261927X18808562.

Liu, F., T. Rahwan, and B. AlShebli. 2023. "Non-White Scientists Appear on Fewer Editorial Boards, Spend More Time Under Review, and Receive Fewer Citations." *Proceedings of the National Academy of Sciences* 120 (13): e2215324120. https://doi.org/10.1073/pnas.2215324120.

Luo, J., T. Feliciani, M. Reinhart, J. Hartstein, V. Das, O. Alabi, and K. Shankar. 2021. "Analyzing Sentiments in Peer Review Reports: Evidence from Two Science Funding Agencies." *Quantitative Science Studies* 2 (4): 1271–1295. https://doi.org/10.1162/qss_a_00156.

Malički, M., and B. Mehmani. 2024. "Structured Peer Review: Pilot Results from 23 Elsevier Journals." *PeerJ* 12:e17514. https://doi.org/10.7717/peerj.17514.

Merton, R. K. 1968. "The Matthew Effect in Science." *Science* 159 (3810): 56–63. https://doi.org/10.1126/science.159.3810.56.

Miller, A. R., and C. Segal. 2019. "Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence." *Review of Economic Studies* 86 (5): 2220–2247. https://doi.org/10.1093/restud/rdy051.

Moss-Racusin, C. A., J. F. Dovidio, V. L. Brescoll, M. J. Graham, and J. Handelsman. 2012. "Science faculty's Subtle Gender Biases Favor Male Students." *Proceedings of the National Academy of Sciences* 109 (41): 16474–16479. https://doi.org/10.1073/pnas.1211286109.

Munz, K. P., M. H. Jung, and A. L. Alter. 2020. "Name Similarity Encourages Generosity: A Field Experiment in Email Personalization." *Marketing Science* 39 (6): 1071–1091. https://doi.org/10.1287/mksc.2019.1220.

Murray, D., K. Siler, V. Lariviére, W. Chan, A. Collings, J. Raymond, and C. Sugimoto. 2018. "Gender and International Diversity Improves Equity in Peer Review." *bioRxiv*. https://doi.org/10.1101/400515.

Nguyen, N. T., M. T. Basuray, W. P. Smith, D. Kopka, and D. McCulloh. 2008. "Moral Issues and Gender Differences in Ethical Judgment Using Reidenbach and Robin's (1990) Multidimensional Ethics Scale: Implications in Teaching of Business Ethics." *Journal of Business Ethics* 77 (4): 417–430. https://doi.org/10.1007/s10551-007-9357-9.

Ni, C., E. Smith, H. Yuan, V. Larivière, and C. R. Sugimoto. 2021. "The Gendered Nature of Authorship." *Science Advances* 7 (36): eabe4639. https://doi.org/10.1126/sciadv.abe4639.

Okike, K., K. T. Hug, M. S. Kocher, and S. S. Leopold. 2016. "Single-Blind Vs Double-Blind Peer Review in the Setting of Author Prestige." *JAMA* 316 (12): 1315–1316. https://doi.org/10.1001/jama.2016.11014.

Oslock, W. M., S. S. Lansing, L. R. Coleman, A. G. Oslock, T. M. Pawlik, S. Noria, and S. Husain. 2024. "Gender Bias in Colorectal Surgery Fellowship Letters of Recommendation." *American Journal of Surgery* 227:198–203. https://doi.org/10.1016/j.amjsurg.2023.10.019.

Parsons, C. E., and R. B. Baglini. 2021. "Peer Review: The Case for Neutral Language." *Trends in Cognitive Sciences* 25 (8): 639–641. https://doi.org/10.1016/j.tics.2021.05.003.

Petruzzelli, A. M. 2011. "The Impact of Technological Relatedness, Prior Ties, and Geographical Distance on University–Industry Collaborations: A Joint-Patent Analysis." *Technovation* 31 (7): 309–319. https://doi.org/10.1016/j.technovation.2011.01.008.

Pinho-Gomes, A.-C., A. Vassallo, K. Thompson, K. Womersley, R. Norton, and M. Woodward. 2021. "Representation of Women Among Editors in Chief of Leading Medical Journals." *JAMA Network Open* 4 (9): e2123026. https://doi.org/10.1001/jamanetworkopen.2021.23026.

Popper, K. 1961. *The Logic of Scientific Discovery*. London: Routledge.

Qiu, Z., R. Marszalek, S. M. Man, and G. Li. 2022. "Peer Review Week 2022: An Interview with Rafal Marszalek, Si Ming Man and Guideng Li About Their Views on Research Integrity as Scientists and Editors." *Communications Biology* 5 (1): 1–3. https://doi.org/10.1038/s42003-022-03957-x.

Rai, A. 2016. "Writing a Virtuous Review." *MIS Quarterly* 40 (3): iii–x.

Rao, R. T., and B. Bareham. 2022. "Regression Towards the Mean—A Plea for Civility in Peer Review." *BMJ* 379:o2886. https://doi.org/10.1136/bmj.o2886.

Resnik, D. B., and S. A. Elmore. 2016. "Ensuring the Quality, Fairness, and Integrity of Journal Peer Review: A Possible Role of Editors." *Science and Engineering Ethics* 22 (1): 169–188. https://doi.org/10.1007/s11948-015-9625-5.

Roter, D. L., and J. A. Hall. 2004. "Physician Gender and Patient-Centered Communication: A Critical Review of Empirical Research." *Annual Review of Public Health* 25 (1): 497–519. https://doi.org/10.1146/annurev.publhealth.25.101802.123134.

Rubin, A., E. Rubin, and D. Segal. 2023. "Editor Home Bias?" *Research Policy* 52 (6): 104766. https://doi.org/10.1016/j.respol.2023.104766.

Saunders, C. 2005. "Looking for Diamond Cutters." *MIS Quarterly* 29 (1): iii–viii. https://doi.org/10.2307/25148664.

Saxenian, A. 2002. "Silicon Valley's New Immigrant High-Growth Entrepreneurs." *Economic Development Quarterly - ECON DEV Q* 16 (1): 20–31. https://doi.org/10.1177/0891242402016001003.

Sikdar, S., M. Marsili, N. Ganguly, and A. Mukherjee. 2017. "Influence of Reviewer Interaction Network on Long-Term Citations: A Case Study of the Scientific Peer-Review System of the Journal of High Energy Physics." *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 1–10. https://doi.org/10.1109/JCDL.2017.7991572.

Skjeie, H. 1991. "The Rhetoric of Difference: On Women's Inclusion into Political Elites." *Politics & Society* 19 (2): 233–263. https://doi.org/10.1177/003232929101900205.

Smith, O. M., K. L. Davis, R. B. Pizza, R. Waterman, K. C. Dobson, B. Foster, J. C. Jarvey, L. N. Jones, W. Leuenberger, N. Nourn, et al. 2023. "Peer Review Perpetuates Barriers for Historically Excluded Groups." *Nature Ecology & Evolution* 7 (4), 512–523. https://doi.org/10.1038/s41559-023-01999-w.

So, C. Y. K. 1998. "Citation Ranking versus Expert Judgment in Evaluating Communication Scholars: Effects of Research Specialty Size and Individual Prominence." *Scientometrics* 41 (3): 325–333. https://doi.org/10.1007/BF02459049.

Straub, W. D. 2008. "Type II Reviewing Errors and the Search for Exciting Papers." *MIS Quarterly* 32 (2): v–x. https://doi.org/10.2307/25148837.

Sun, Z., C. Clark Cao, C. Ma, and Y. Li. 2023. "The Academic Status of Reviewers Predicts Their Language Use." *Journal of Informetrics* 17 (4): 101449. https://doi.org/10.1016/j.joi.2023.101449.

Tang, L., L. Wang, and G. Hu. 2023. "Research Misconduct Investigations in China's Science Funding System." *Science and Engineering Ethics* 29 (6): 39. https://doi.org/10.1007/s11948-023-00459-9.

Teplitskiy, M., D. Acuna, A. Elamrani-Raoult, K. Körding, and J. Evans. 2018. "The Sociology of Scientific Validity: How Professional Networks Shape Judgement in Peer Review." *Research Policy* 47 (9): 1825–1841. https://doi.org/10.1016/j.respol.2018.06.014.

Tsugawa, Y., A. B. Jena, J. F. Figueroa, E. J. Orav, D. M. Blumenthal, and A. K. Jha. 2017. "Comparison of Hospital Mortality and Readmission Rates for Medicare Patients Treated by Male Vs Female Physicians." *JAMA Internal Medicine* 177 (2): 206–213. https://doi.org/10.1001/jamainternmed.2016.7875.

Vermeulen, F., and H. Barkema. 2001. "Learning Through Acquisitions." *Academy of Management Journal* 44 (3): 457–476. https://doi.org/10.2307/3069364.

Wallis, C. J. D., A. Jerath, K. Aminoltejari, K. Kaneshwaran, A. Salles, N. Coburn, F. C. Wright, L. Gotlib Conn, Z. Klaassen, A. N. Luckenbaugh, et al. 2023. "Surgeon Sex and Long-Term Postoperative Outcomes Among Patients Undergoing Common Surgeries." *JAMA Surgery* 158 (11): 1185. https://doi.org/10.1001/jamasurg.2023.3744.

Wallis, C. J. D., A. Jerath, N. Coburn, Z. Klaassen, A. N. Luckenbaugh, D. E. Magee, A. E. Hird, K. Armstrong, B. Ravi, N. F. Esnaola, et al. 2022. "Association of Surgeon-Patient Sex Concordance with Postoperative Outcomes." *JAMA Surgery* 157 (2): 146–156. https://doi.org/10.1001/jamasurg.2021.6339.

Wing, D. A., R. S. Benner, R. Petersen, R. Newcomb, and J. R. Scott. 2010. "Differences in Editorial Board Reviewer Behavior Based on Gender." *Journal of Women's Health* 19 (10): 1919–1923. https://doi.org/10.1089/jwh.2009.1904.

Witteman, H. O., M. Hendricks, S. Straus, and C. Tannenbaum. 2019. "Are Gender Gaps Due to Evaluations of the Applicant or the Science? A Natural Experiment at a National Funding Agency." *Lancet* 393 (10171): 531–540. https://doi.org/10.1016/S0140-6736(18)32611-4.

Wolfram, D., P. Wang, and F. Abuzahra. 2021. "An Exploration of referees' Comments Published in Open Peer Review Journals: The Characteristics of Review Language and the Association Between Review Scrutiny and Citations." *Research Evaluation* 30 (3): 314–322. https://doi.org/10.1093/reseval/rvab005.

Xu, S., G. Zhang, Y. Sun, and X. Wang. 2019. "Understanding the Peer Review Endeavor." *Proceedings of the Association for Information Science and Technology* 56 (1): 316–325. https://doi.org/10.1002/pra2.26.

Zhang, G., F. Shang, L. Wang, W. Xie, P. Jia, C. Jiang, and X. Wang. 2023. "Is Peer Review Duration Shorter for Attractive Manuscripts?" *Journal of Information Science* 016555152311743. https://doi.org/10.1177/01655515231174382.

Zhang, G., L. Wang, W. Xie, F. Shang, X. Xia, C. Jiang, and X. Wang. 2022. "'This Article is Interesting, However': Exploring the Language Use in the Peer Review Comment of Articles Published in the BMJ." *Aslib Journal of Information Management* 74 (3): 399–416. https://doi.org/10.1108/AJIM-06-2021-0172.

Zhang, G., S. Xu, Y. Sun, C. Jiang, and X. Wang. 2022. "Understanding the Peer Review Endeavor in Scientific Publishing." *Journal of Informetrics* 16 (2): 101264. https://doi.org/10.1016/j.joi.2022.101264.

Zheng, X., J. Chen, A. Tollas, and C. Ni. 2023. "The Effectiveness of Peer Review in Identifying Issues Leading to Retractions." *Journal of Informetrics* 17 (3): 101423. https://doi.org/10.1016/j.joi.2023.101423.

Zmud, B. 1998. "A Personal Perspective on the State of Journal Refereeing." *MIS Quarterly* 22 (3): xlv–xlviii.