Descriptive statistics

Adrian Barnett

05 July, 2025

These results are for the analysis data set where the reviewer could be matched.

## Summary table

level Overall   
 n 37332   
 year (median [IQR]) 2022.0 [2019.0, 2024.0]  
 journal (%) F1000Research 24132 (64.6)   
 Gates Open Research 1714 ( 4.6)   
 Open Research Europe 2789 ( 7.5)   
 Wellcome Open Research 8697 (23.3)   
 role (%) Co-referee 2428 ( 6.5)   
 Referee 34904 (93.5)   
 recommendation (%) Approve 19984 (53.5)   
 Reservations 14379 (38.5)   
 Reject 2969 ( 8.0)   
 version3 (%) 1 26474 (70.9)   
 2 8995 (24.1)   
 3+ 1863 ( 5.0)   
 n\_papers\_cited (median [IQR]) 24.0 [14.0, 38.0]   
 matches (mean (SD)) 0.2 (0.7)   
 any\_matches (%) No 32375 (86.7)   
 Yes 4957 (13.3)   
 n\_reviewer\_cited (mean (SD)) 0.4 (1.5)   
 any\_cited (%) No 31546 (84.5)   
 Yes 5786 (15.5)   
 self\_cited\_count (mean (SD)) 0.1 (0.5)   
 any\_self (%) No 35023 (93.8)   
 Yes 2309 ( 6.2)   
 works\_count (median [IQR]) 55.0 [24.0, 118.0]   
 n\_words (median [IQR]) 202.0 [67.0, 411.0]

## Top ten reviewers’ countries

| **country** | **n (%)** |
| --- | --- |
| USA | 7655 (21%) |
| United Kingdom | 4137 (11%) |
| India | 2472 (7%) |
| Italy | 1368 (4%) |
| Australia | 1349 (4%) |
| Germany | 1277 (3%) |
| China | 1235 (3%) |
| Canada | 1141 (3%) |
| Indonesia | 1108 (3%) |
| Spain | 858 (2%) |

There were 0 reviews where the country was missing.

## Number of reviewers per article

| **n** | **Q1** | **median** | **Q3** |
| --- | --- | --- | --- |
| 18,169 | 1 | 2 | 2 |

The summary statistics are for the number of reviewers per article and version.

## Citation numbers by reviewers’ experience

This is an internal validation check of the data by examining if the number of times a reviewer is cited in the article increases with their experience as measured by their count of published papers. We used a Poisson regression model with a dependent variable of the count of article citations and independent variable of the reviewers’ publication counts.

| **term** | **estimate** | **conf.low** | **conf.high** | **p.value** |
| --- | --- | --- | --- | --- |
| log(reviewer`s publication count) | 38.3 | 36.4 | 40.1 | < 0.001 |

We used the log base 2 transform of the reviewer’s publication counts, so the estimates show the percentage change in the number of citations per doubling of the reviewer’s publication counts. As expected, there’s a strong increase in the number of citations for more experienced reviewers.

## Citation probability by reference list size

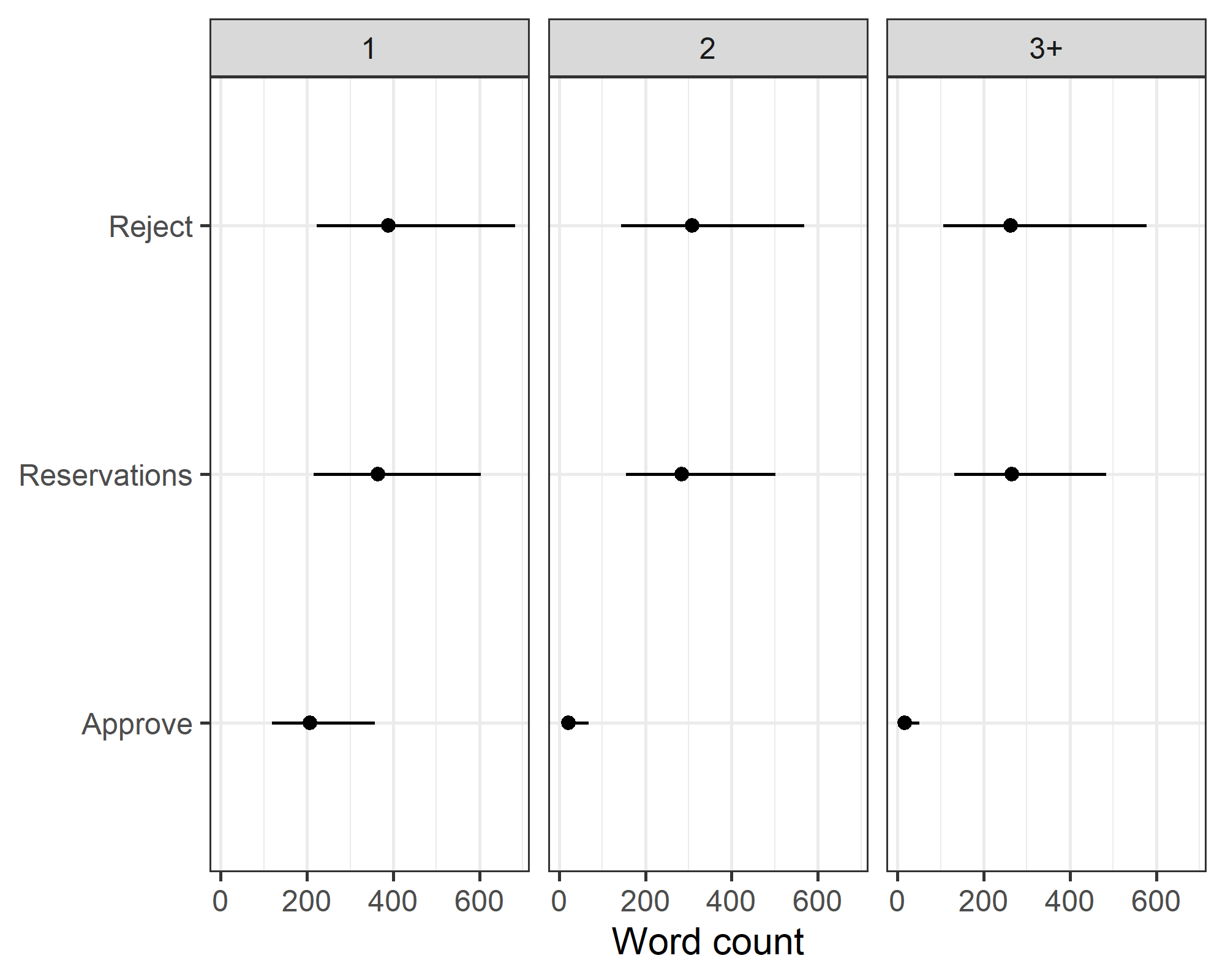
This is an internal validation check of the data by examining if the number of citations to the reviewer increases with the total number of papers cited in the article. We used a Poisson regression model with a dependent variable of the count of article citations and independent variable of the total number of citations in the article.

| **term** | **estimate** | **conf.low** | **conf.high** | **p.value** |
| --- | --- | --- | --- | --- |
| log(reference list size) | 70.2 | 66.6 | 73.8 | < 0.001 |

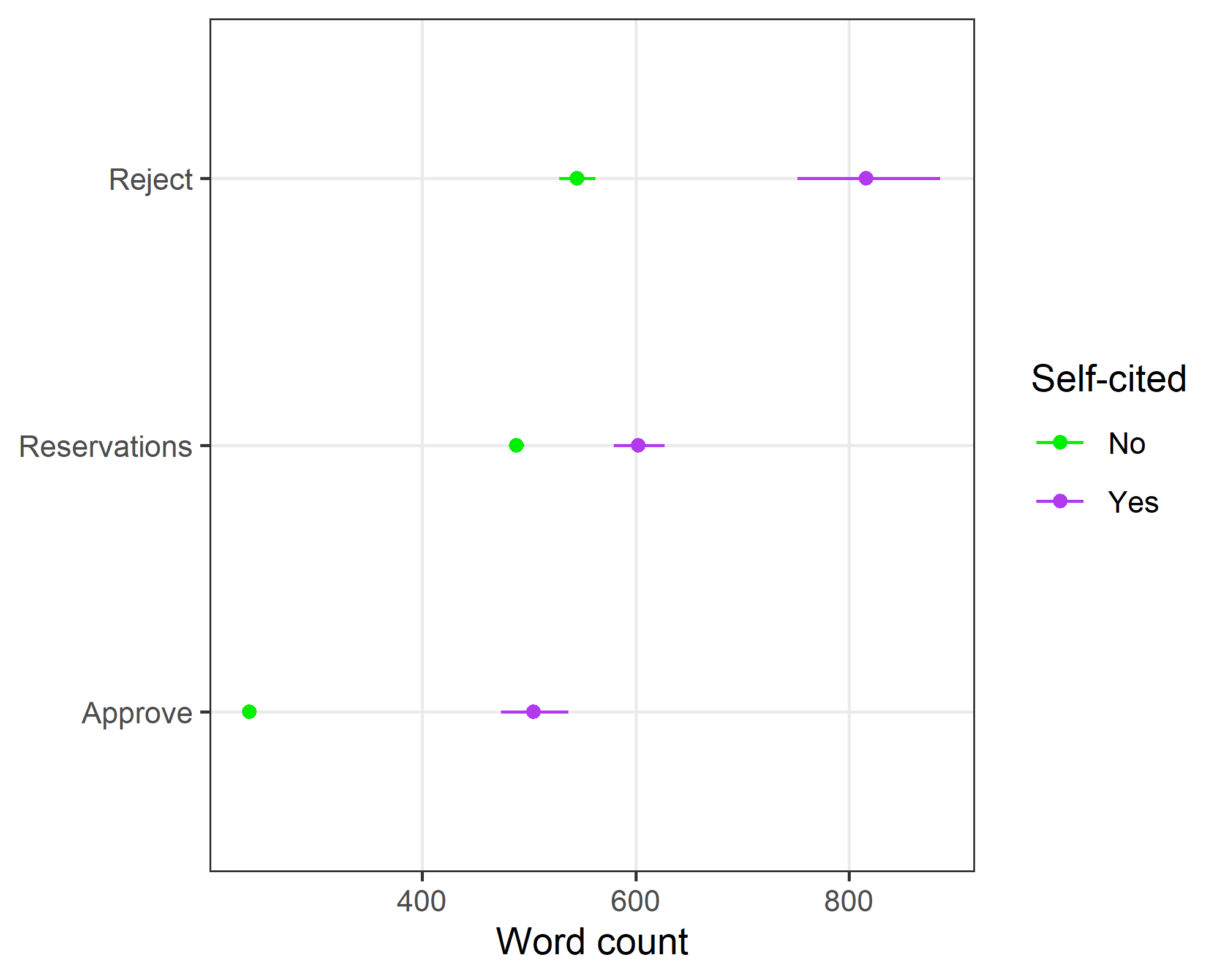
We used the log base 2 transform of the reference list, so the estimates show the percentage change in the number of citations per doubling of the number of papers in the article’s reference list. As expected, there’s a strong increase in the number of citations to the reviewer when the article has cited more papers.

## Review length

Here we plot the length of the review by version and recommendation.



### Review length by self-citations



#### Table of estimates

| **Term** | **estimate** | **conf.low** | **conf.high** |
| --- | --- | --- | --- |
| Version > 1 | 0.51 | 0.49 | 0.52 |
| Reservations | 2.05 | 2.00 | 2.10 |
| Reject | 2.29 | 2.21 | 2.37 |
| Self-cited | 2.12 | 1.98 | 2.26 |
| Reservations x self-cited | 0.58 | 0.54 | 0.63 |
| Reject x self-cited | 0.71 | 0.63 | 0.79 |