Supplemental Information: How Gendered is the Peer Review Process? A Mixed-Design Analysis of Reviewer Feedback *

Thomas König *University of Mannheim*
Guido Ropers *University of Mannheim*

Table S1: Review score estimation for all possible combinations for a two-recommendation set following Bravo et al. (2018, 104).

Recommendations	Potential recommendation set			Review score
	#better	#worse	#unclear	(optimistic)
{accept, accept}	0	9	0	1.00
{accept, minor revisions}	1	8	0	0.89
{accept, major revisions}	2	6	1	0.75
{minor revisions, minor revisions}	2	5	2	0.71
{minor revisions, major revisions}	4	4	1	0.50
{accept, reject}	3	3	3	0.50
{major revisions, major revisions}	5	2	2	0.29
{minor revisions, reject}	6	2	1	0.25
{major revisions, reject}	8	1	0	0.11
{reject, reject}	9	0	0	0.00

*This version: June 24, 2021. Email: tkoenig@uni-mannheim.de and gropers@mail.uni-mannheim.de

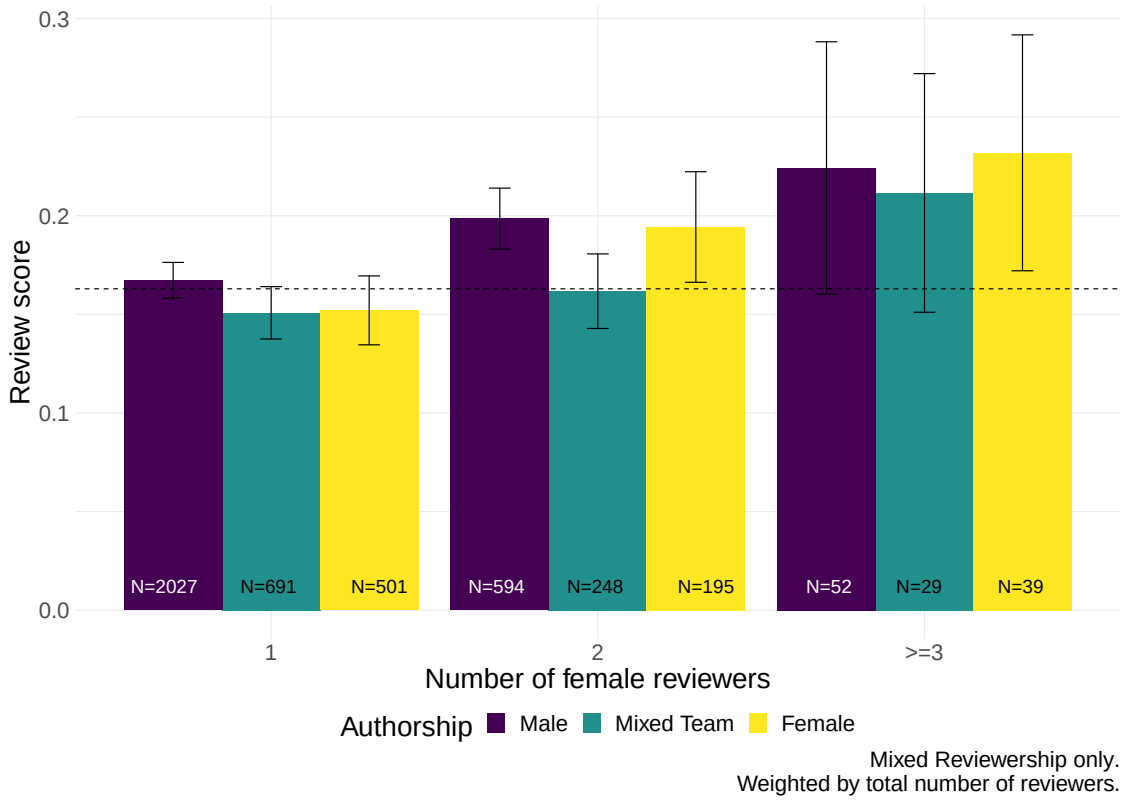

Figure S1: Average review scores for mixed reviewer gender compositions conditional on authors' gender composition. Error bars show 95% confidence intervals. The dotted horizontal line indicates the average review sore in the sample.

Table S2: OLS Regression Results: Tone in review reports conditional on reviewer gender and authorship type

	Negativity I	Negativity II	Positivity I	Positivity II	Disgust I	Disgust II
Female Reviewer	−0.00	−0.01	0.16***	0.14**	−0.00	0.01
	(0.02)	(0.04)	(0.02)	(0.05)	(0.01)	(0.01)
Male * Female Reviewer		0.01		0.03		−0.01
		(0.04)		(0.06)		(0.02)
Female * Female Reviewer		0.00		0.01		−0.02
		(0.05)		(0.08)		(0.02)
Manuscript FE	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	20267	20267	20267	20267	20267	20267
Num. Manuscripts	7069	7069	7069	7069	7069	7069

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Outcome variables are the share of negatively associated words, the share of positively associated words, and the share of words associated with disgust. Reference group are male reviewers and mixed team authorship. Standard errors are clustered at the manuscript level.

Table S3: OLS regression results: Gender reviewing behavior conditional on disaggregated authorship type (solo vs. team submissions)

	Non-reject recommendation	Length	Sentiment	Duration
Female Reviewer	0.03^{*}	-15.28	0.16^{**}	2.52^{***}
	(0.01)	(17.46)	(0.05)	(0.64)
Solo Female * Female Reviewer	0.01	-25.98	0.06	-2.01
	(0.03)	(32.47)	(0.10)	(1.20)
Male Team * Female Reviewer	0.00	-18.47	0.02	-1.61
	(0.02)	(25.84)	(0.08)	(0.94)
Female Team * Female Reviewer	0.06	11.49	-0.19	-1.01
	(0.05)	(51.62)	(0.15)	(1.95)
Mixed Gender * Female Reviewer	0.01	27.14	-0.01	-0.72
	(0.02)	(27.66)	(0.08)	(1.01)
Manuscript FE	Yes	Yes	Yes	Yes
Num. obs.	23948	20267	20267	23948
Num. manuscripts	8294	7069	7069	8294

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Outcome variables are non-reject recommendations, the number of words in a review ("Length"), a sentiment score measured as difference in the percentage share of positive and negative words in a review ("Sentiment"), and the duration (days) until a review is being submitted ("Duration"). Reference group are male reviewers and solo male authors. Standard errors are clustered at the manuscript level.

Table S4: OLS Regression Results

	Review Score	Non-Reject Decision
Intercept	0.14^{***}	-0.05^{***}
	(0.01)	(0.01)
Two female reviewer	0.04^{***}	
	(0.01)	
More than two female reviewer	0.08^{***}	
	(0.02)	
Mixed team	-0.01	-0.01
	(0.01)	(0.01)
Female authorship	-0.02	0.01
	(0.01)	(0.01)
Comparative politics	0.01	0.03^{*}
	(0.01)	(0.01)
Formal Theory	0.02	0.07^{**}
	(0.01)	(0.02)
International Relations	-0.02^{*}	-0.01
	(0.01)	(0.01)
Methodology	0.01	0.03
	(0.02)	(0.03)
Normative Theory	0.08^{***}	-0.03^{*}
	(0.01)	(0.01)
Race and Ethnicity	-0.01	0.01
	(0.01)	(0.02)
Other	-0.03	-0.00
	(0.02)	(0.02)
Review Score		1.02^{***}
		(0.02)
R^2	0.04	0.34
Adj. R^2	0.04	0.34
Num. obs.	4376	4376

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. The sample consists of manuscripts with mixed reviewership. Reference group are manuscripts with one female reviewers and male authorship.

References

Bravo, Giangiacomo et al. (2018). "Hidden Connections: Network Effects on Editorial Decisions in Four Computer Science Journals". In: *Journal of Informetrics* 12.1, pp. 101–112.