Supplementary materials: The impact of double blind reviewing at EvoLang 11

1 Assessing the gender-typing of EvoLang

Previous studies showed that the bias against females is stronger in topics which are perceived to be more masculine (Knobloch-Westerwick et al., 2013). No independent measure of topic gender typing was available, so the rates of submissions by male and female authors for each topics were used as a proxy. The prediction would be that double blind review would increase the ratings of female papers, but less so for more female-type topics.

In EvoLang 11 each corresponding author categorised their paper as belonging to at least one of 21 topics. The proportion of male and female first authored papers for each topic was calculated, disregarding topics for which there were fewer than 5 papers. Table 1 shows the results, ranked from most male-dominated to most female dominated. Subjectively, the ordering makes sense: artificial life is the most male-type topics while gesture and sign language are the most female-type topics.

A given paper was assigned a male topic type score by taking the mean proportion of male submissions for each topic assigned to that paper. Individual authors were then given a male topic type score by taking the mean score for each of their papers. That is, each author was assigned a score which reflected how gender biased the submissions for their topics were.

The paired change in ranking data was analysed (since topics were not assigned in EvoLang 10 papers). Male topic type score showed a non-significant negative correlation with the change in paper ranking (the more male-biased the topic, the smaller the improvement in ranking, r = -0.2, p = 0.09). The difference in the maximum paper score for each author between the two conferences was entered as the dependent variable in a linear model, with gender, male topic type score and the interaction between the two as independent variables. There were no significant effects (gender: t = -0.75, 0 = 0.16; male topic type score: -1.47, p = 0.15; interaction: t = 0.71, p = 0.48).

That is, there was no detectable effect of topic gender type on the results.

Table 1: Percentage of female and male first authors by topic.

Topic	Female first authors (%)	Male first authors (%)
artificial life	8	92
genetics	15	85
anthropology	19	81
modeling	19	81
phonology	21	79
lexicon	24	76
phonetics	25	75
syntax	28	72
biology	31	69
semantics	35	65
neuroscience	35	65
acquisition	36	64
physiology	36	64
cognitive science	38	62
pragmatics	40	60
psychology	40	60
primatology	45	55
sign language	65	35
gesture	70	30

2   Country of affiliation

The number of authors, submitted papers and accepted papers by country are shown in table 2 for all authors in EvoLang 11. Within the represented countries, there is no obvious bias. For example, the number of submissions from a country does not predict the likelihood of acceptance (ρ = -0.22, p = 0.28). However, only Europe and the USA are well represented, while many parts of the world represented poorly or not at all.

References

Knobloch-Westerwick, S., Glynn, C. J., and Huge, M. (2013). The matilda effect in science communication an experiment on gender bias in publication quality perceptions and collaboration interest. *Science Communication*, 35(5):603–625.

Table 2: Submitted and accepted papers by country of affiliation.

	Submitting Authors	Submitted Papers	Accepted Papers
Brazil	1	1	1
Denmark	1	1	1
Hungary	1	1	1
Ireland	1	1	0
Philippines	1	1	0
Estonia	2	2	1
Lebanon	2	2	2
Norway	2	1	1
Poland	4	2	2
Singapore	4	2	1
Israel	6	3	3
Sweden	7	5	3
Switzerland	8	5	5
Spain	9	6	6
Canada	10	7	6
Italy	12	6	6
Australia	13	7	6
France	20	6	5
Austria	21	10	8
Belgium	27	13	12
Japan	29	18	6
Netherlands	36	19	18
Germany	50	26	20
United Kingdom	113	48	40
United States	113	54	43