Language use shapes cultural norms: Large scale evidence from gender

Molly Lewis[1,2] and Gary Lupyan[1]

[1]University of Wisconsin-Madison, [2]University of Chicago

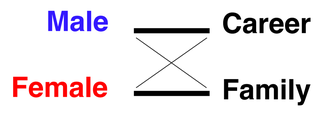
What role do word co-occurrences and grammatical structure play in shaping cultural norms?

Gender bias as a case study: an abstract domain,[1] that is culturally transmitted and often grammatically encoded.[2]

Hypotheses: (1) Language as a mere reflection of speakers' gender biases, or (2) language as one of the causes?

Study 1: Gender bias across cultures

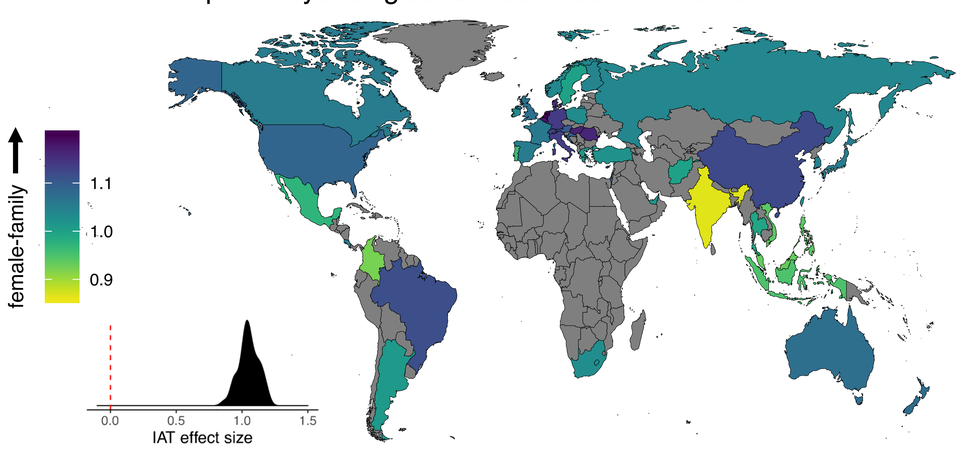
Implicit Association Task (IAT) – behavioral measure of the strength of respondents' implicit associations between two pairs of concepts.[3]



Data collected by Project Implicit[4] – 663,709 participants from 48 countries (d = 1.08; M = 1.05; SD = .07).

Implicit and explicit bias measures correlated (r = .15; p < .0001).

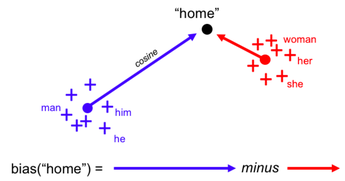
Implicit Psychological Gender Bias Effect Size



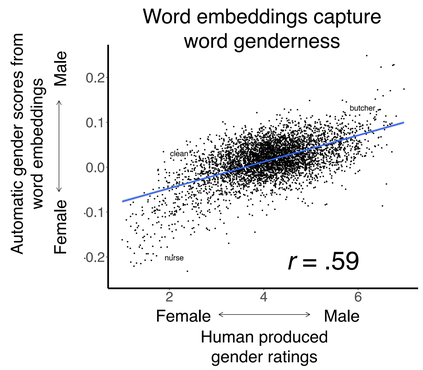
References: **[1]** Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, 43(1), 1-22.; **[2]** Master, A., Markman, E. M., & Dweck, C. S. (2012). Thinking in categories or along a continuum: Consequences for children's social judgments. *Child Development*, 83(4), 1145-1163.; **[3]** Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *JPSP*, 74(6), 1464.; **[4]** Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, 6(1), 101.; **[5]** Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2016). Enriching word vectors with subword information.; **[6]** Scott, G. G., Keitel, A., Becirspahic, M., O'Donnell, P. J., & Sereno, S. C. (2017). The Glasgow Norms: Ratings of 5,500 words on 9 scales. **[7]** Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183-186.

Study 2: Gender bias and word meanings

We quantified gender bias *in language* using distributional semantic models (word-embeddings trained on Wikipedia.[5])

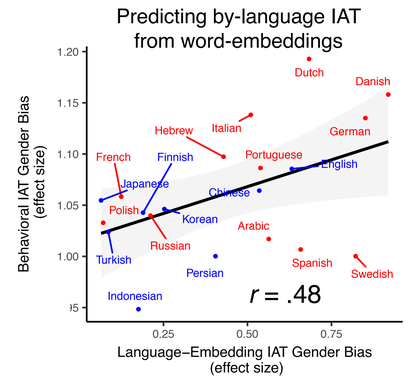


In English, gender bias of a word in embedding model correlated strongly with explicit gender ratings[6] (r = .59; p <. 001).



We then computed the linguistic analog to behavioral IAT using English word embedding models (replicating Caliskan, et al., 2017[7]) when trained on corpora from each of the languages. Target words were translated into 20 languages by native speakers.

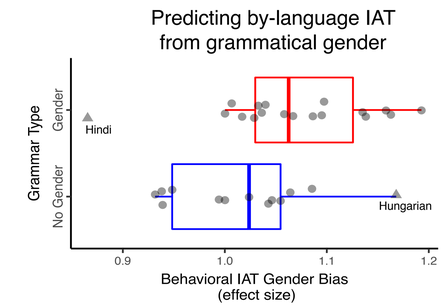
Behavioral and language IAT measures positively correlated at the level of languages (r = .48; p = .03).



Study 3: Gender bias and grammar

Grammatically gendered languages show somewhat higher language gender biases, compared to languages without grammatical gender (d = .99 [-.02, 2.01]).

Participants who completed IAT in countries where the dominant language has grammatical gender showed a somewhat larger gender bias on the IAT (d = 0.68 [-0.08, 1.45]).



Conclusions

People exposed to languages that encode stronger gender biases (Study 2) and those with grammatical gender (Study 3) tend to show larger implicit gender biases on the (English) IAT.

A relationship between grammatical gender and IAT hints at a causal influence: input from language structure affects IAT performance.