

POWERFUL

CREATIVE

PRETTY

LOGICAL

nature  
human behaviour

ARTICLES

<https://doi.org/10.1038/s41562-020-0918-6>

 Check for updates

# Gender stereotypes are reflected in the distributional structure of 25 languages

Molly Lewis <sup>1,2</sup>  and Gary Lupyan <sup>3</sup>



# 01

问题的提出

Gender stereotypes

---



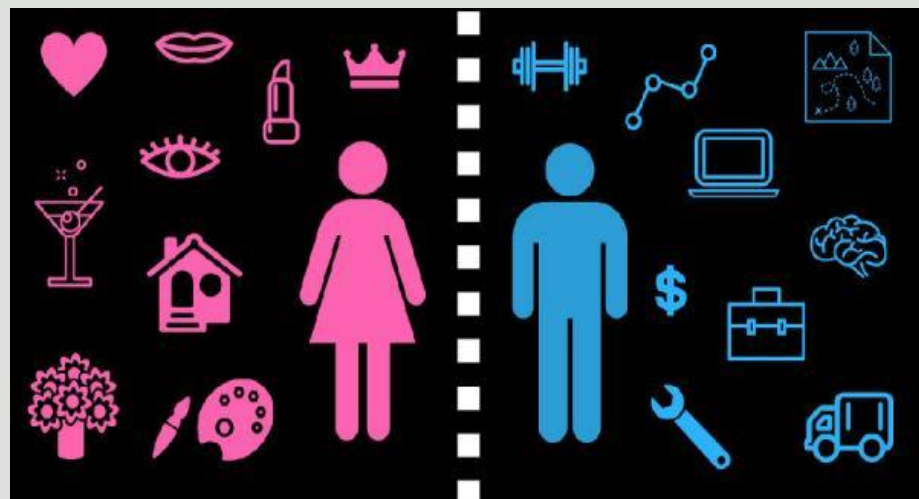
## 性别刻板印象

男性更擅长工作？  
女性则擅长照顾家庭？

这些刻板印象在语言结构中是否能反映？

女孩不太可能被称赞很聪明，也不会认为自己很聪明；男孩子则从小就被鼓励。

这些信念上的差异来自哪里？



# 语言 & 刻板印象

## 我们为什么会有刻板印象？

1. 直接经验
2. 语言

有时两个有着相同直接经验的人，也会因为语言的不同产生不同的刻板映像。

语言与刻板映像之间的对应关系是否在广泛的语言中都存在？



# 02

主要工作

THE MAIN WORK

---



# 做什么？

## ● study 1

基于不同语言训练的模型对性别的联想是否与IAT呼应。



## study 2

心理测量和模型计算的性别联想  
如何与语言结构关联？



# 怎么做？

## 相关工具和实验

### 验证指标

- IAT (the Implicit Association Test)
- 特定职业IAT：男性&职业VS女性&家庭

### 语言模型

- 分布语义模型（共现）
- dog-cay \*banana --->women-home&family

### 实验一

- 1.a 用IAT验证语言模型
- 1.b 在每种语言的文本上训练的词嵌入模型
- 1.c 比较英式英语和美式英语的差异，推广更多主题

### 实验二

- 语法上的性别
- 现存的职业性别术语



# 03

方法

## METHOD

---





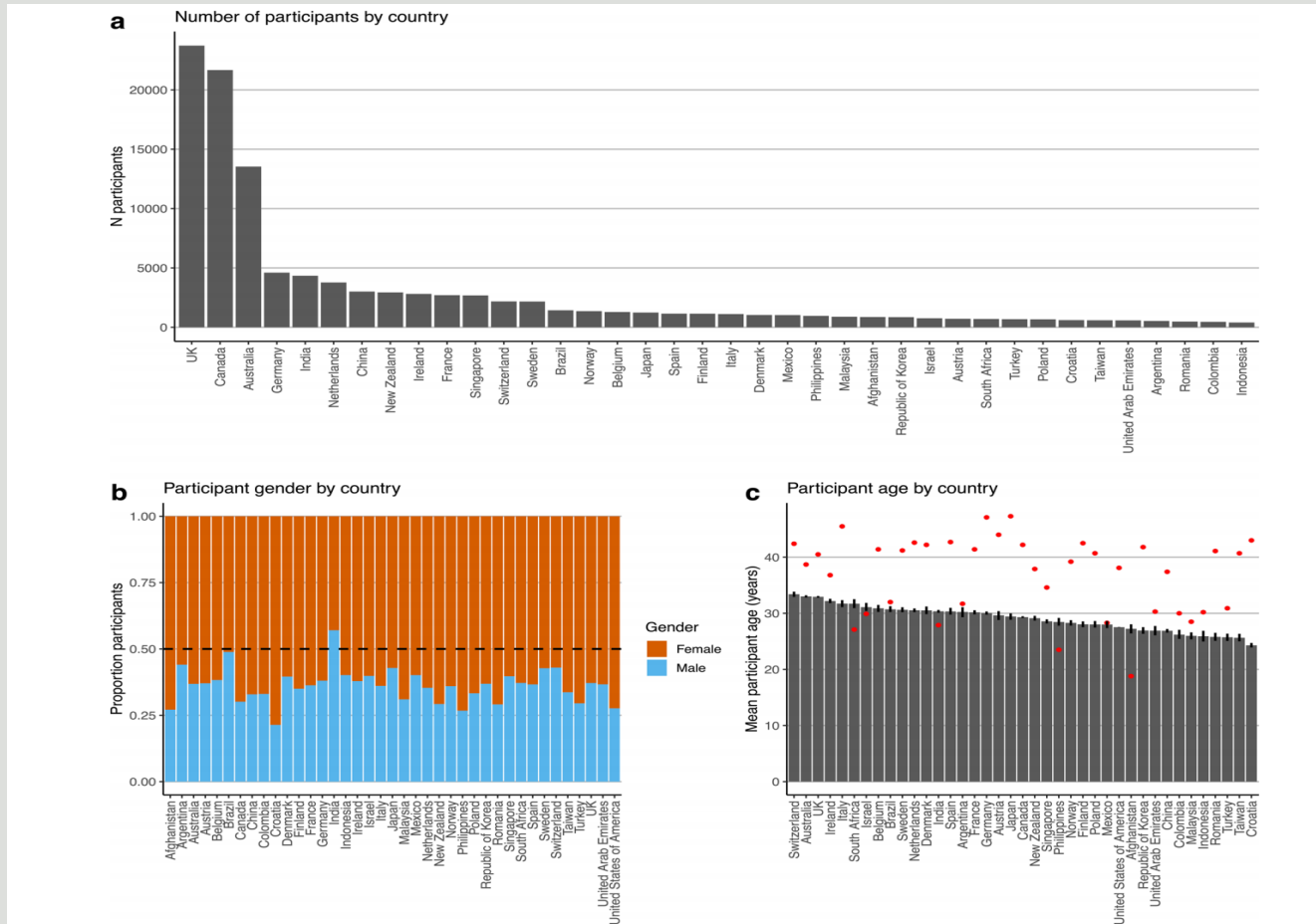
pearson  $r$ ;

two sample t-test : welch' s test,

Effect size : classic Cohen' s  $d$ ;

95% confidence intervals,two-sided analyses.





样本

657,335 participants from 39 countries, with a median of 1,145 participants per country; more female participants; Red points show median age by country.





## 什么是IAT

**male/career , male/family**

**words denoting more similar meanings are easier to pair together** compared to words denoting more dissimilar pairs.

‘How strongly do you associate the following with males and females?’ **from female (1) to male (7)(? ? )**

Project Implicit between 2005-2016; complete demographic information >400; English



## 与前面的结论一致

participants tended to implicitly associate **men with career and women with family**

## 年龄越大联想越强

**Older participants showed greater implicit associations** between women-family and men-career.

## 女性参与者联想更强

The measured associations were **stronger for female participants**

## 不一致的刻板印象反映更强

participants that **received the block of trials with stereotype-incongruent mappings first** than those who received the stereotype-incongruent mappings second

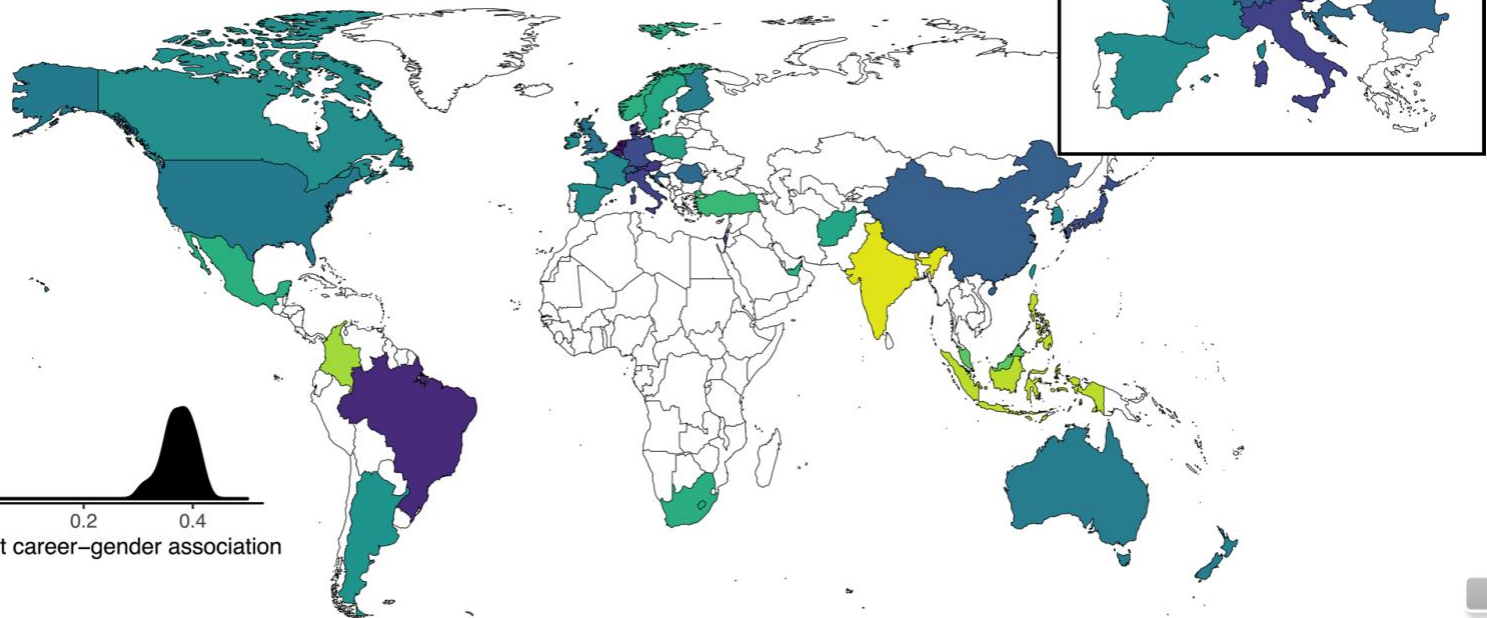
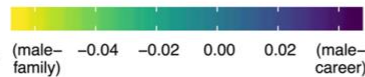


## Geographic distribution of IAT scores

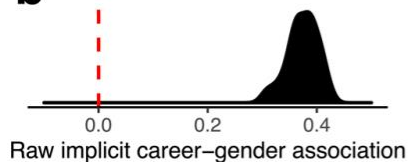
IAT scores are residualized for participant age, gender, and task order (N = 657,335)

**a**

Residualized implicit  
career–gender association



**b**



## 这些数据能反映真实的情况吗？

把每个国家的分数和一份报告 (Scientific and Cultural Organization;2012-2017)做了比较，比较在STEM领域女性参与的百分比。

每个国家的得分和隐形职业性别联想分数越低，女性在科学领域的参与度越高。联想的强弱和每个国家参与者的年龄有关。

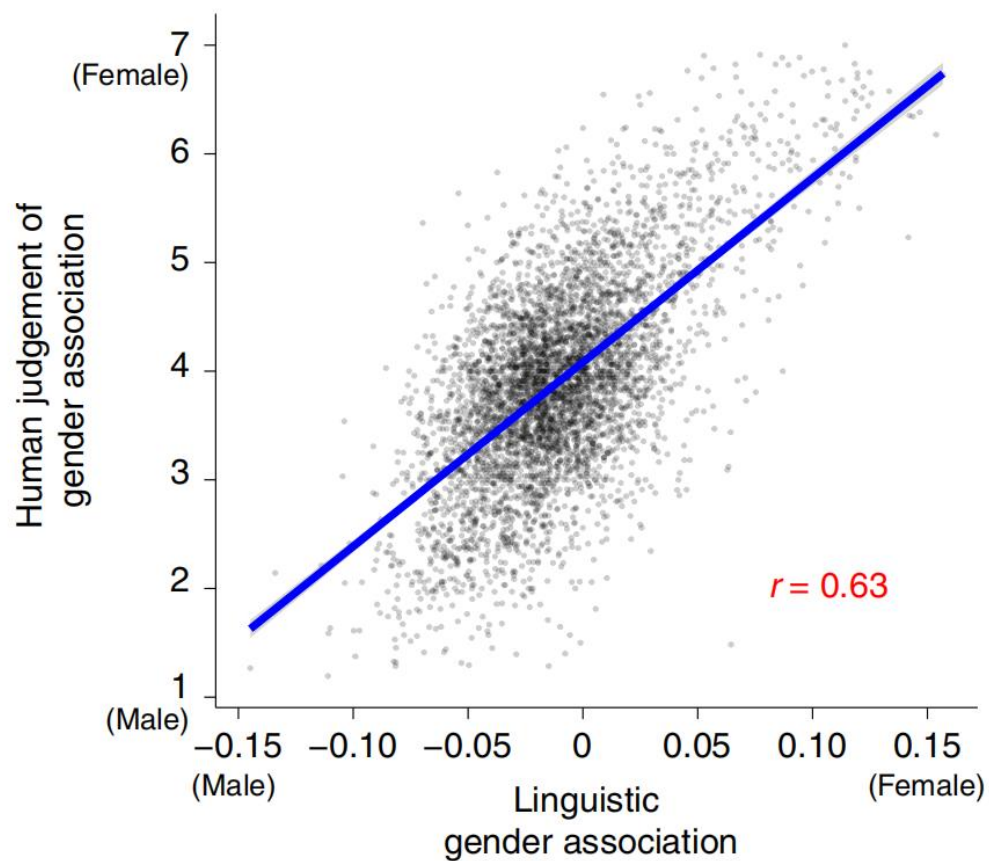


## study1a

1. estimated a gender score for each word by **measuring the average cosine distance to a standard set** of male or female.(male-man/he/boy/son) ,
2. estimated gender scores for each word from models pre-trained on : **subtitles&Wikipedia.[fastText**<sub>(<https://zhuanlan.zhihu.com/p/32965521>)</sub>**]**
3. association estimates from both word-embedding models were **highly correlated with human judgements** of word gender.



## 心理上的性别联想可以通过词嵌入模型测量





1. **estimated** the magnitude of these associations in the **dominant language spoken** in each country represented in the Project Implicit dataset [(Wikipedia=25; subtitle=20)) ]
2. **compare** this with psychological career–gender associations .

most frequently spoken language in each country;

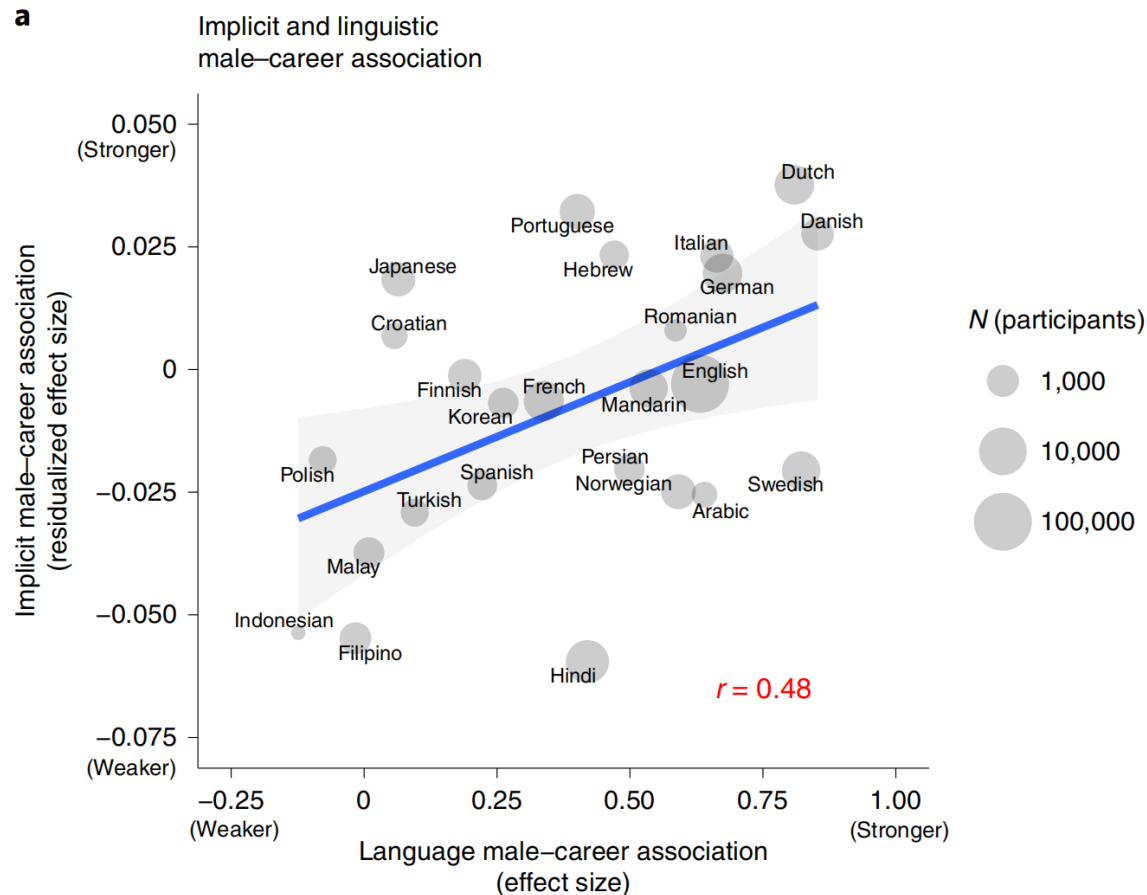
25;

obtained translations;

8 female /male ‘target words’ -eight careers/families



## Implicit VS linguistic associations.

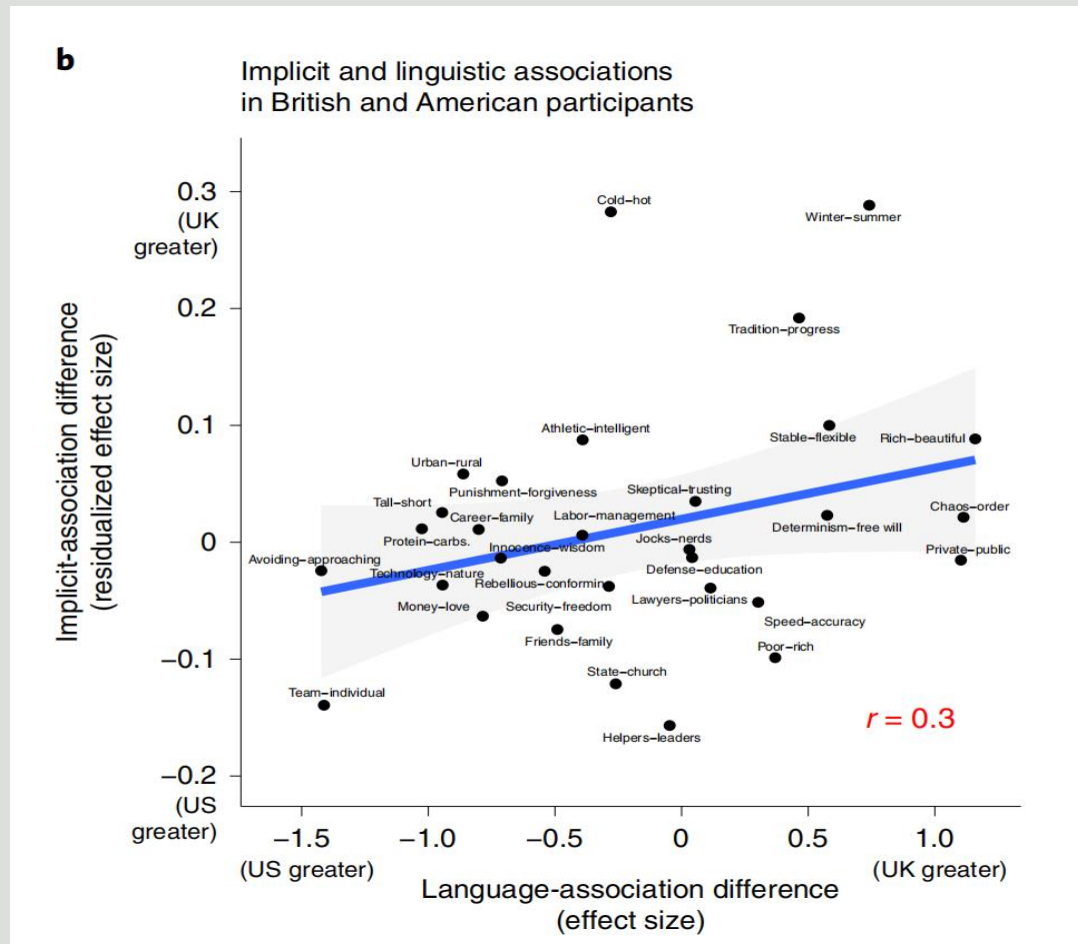


## study1c (confirmatory)

1. Attitudes, Identities and Individual Differences Study dataset (AIID), containing measures of IAT performance from more than 200,000 participants for a wide range of IATs
1. compare two different dialects of English: British[BNC]) and American English[COCA]



the degree to which those associations were present in a speaker's English dialect would predict the magnitude of their psychological association



## 试验一的结论

- 
1. previously reported psychological gender association was correlated with the magnitude of that same association in the language statistics of 25 languages.
  2. This pattern extends to associations beyond career and gender. These results suggest a close correspondence between psychological and linguistic gender associations.





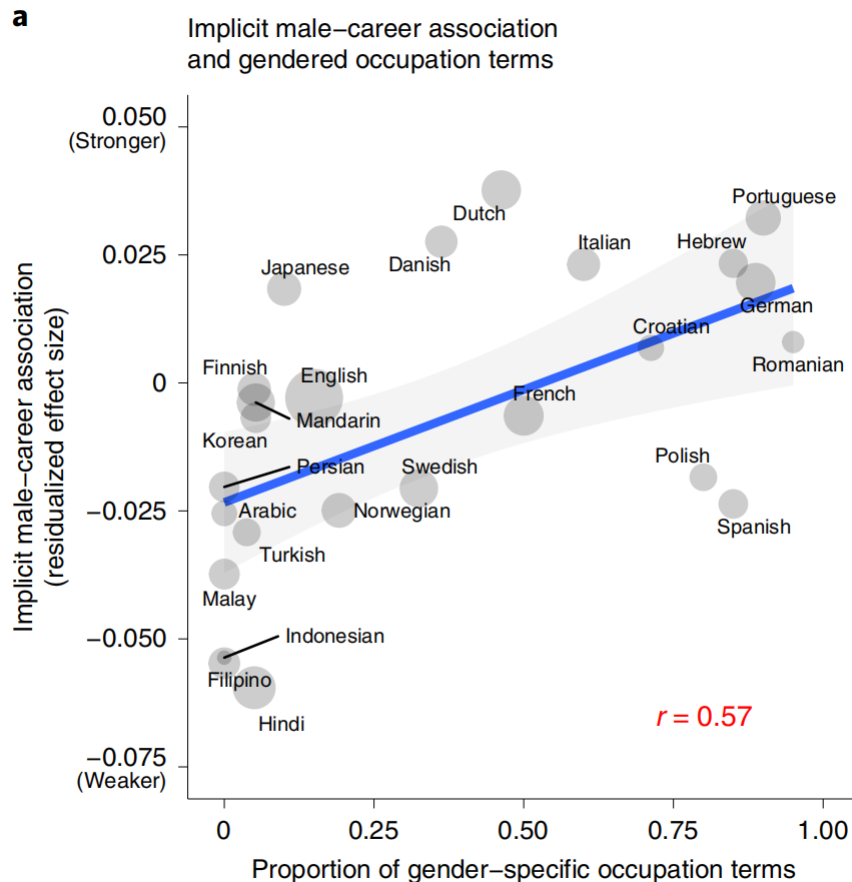
Une bonne étudiante  
Un bon étudiant

waiter  
waitress

stable features of language

- whether grammatical gender and use of gender-specific occupation terms are associated with a greater psychological gender association
- whether this relationship is further mediated by language statistics.



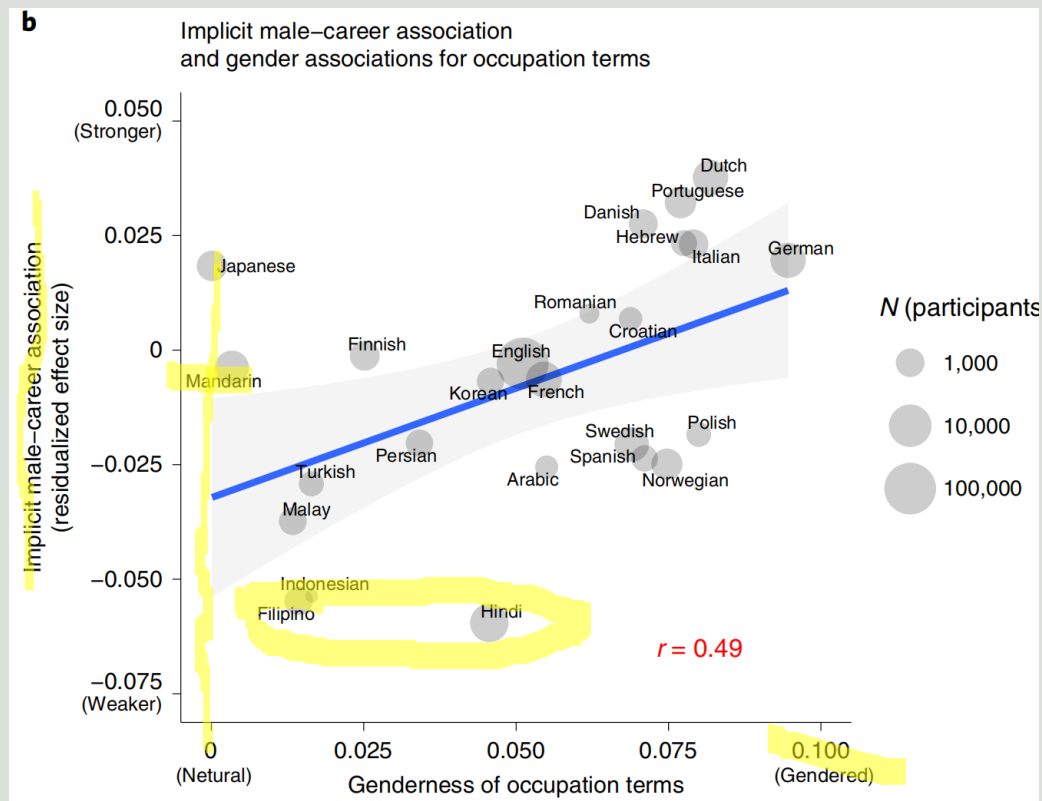


- languages with more gender-specific forms tended to have speakers with greater implicit career–gender association
- There was no relationship between explicit psychological career–gender association and lexical marking of occupation words



mixed-effects model predicting degree of gender association in language statistics from distinctiveness between male and female forms for that word. 'secretary' has a greater gender association in Italian compared with English.

languages with more gendered occupation terms had stronger career–gender associations in their language statistics ( $r = .6$  &  $r = .78$ )





## 试验二的结论

one driver of the relationship between language and psychological career–gender associations observed in study 1 may be the presence of gender–specific occupation terms.



# 04

未来工作

Future Work

---





validity of the IAT. the explicit association measure being too coarse.

?



# 「做什么」



how much exposure to the relevant language statistics is sufficient to produce differences in beliefs

how resilient the learned associations are to other sources of information.(combination or vary dynamically )



POWERFUL

CREATIVE

PRETTY

LOGICAL

# THANK YOU!

