Do speakers consult an internal jury of their peers in judging linguistic 'fault' and subjectivity?

Ruthe Foushee & **Mahesh Srinivasan**
foushee@uchicago.edu srinivasan@berkeley.edu

SPP 2021

Intro

Language is a conventional system[1]

Adults readily make fine-grained judgments of statements' subjectivity, correlated with judgments of disagreements as *faultless* (neither speaker is wrong), and predictive of cross-linguistic phenomena[2-3]

Subjectivity is described as a cognitive universal, but where do adults' graded evaluations of subjectivity *come from*?

We explore the hypothesis that adults' graded evaluations of faultless disagreement/subjectivity/relative truth derive from modeling their own speech community.

Experiment 1 *Are faultless disagreement judgments systematically related to estimates of population-level consensus?*

Method

Participants 204 Amazon's Mechanical Turk Workers

Stimuli 14 t-shirt images, 7 in prototypical hues, 7 borderline (e.g., BLUE-GREEN), according to WCS[4] naming data

Adjectives 7 color terms, 7 evaluative predicates

red, orange, yellow, green, blue, purple, pink
pretty, nice, exciting, pleasant, boring, ugly, strange

2 Blocks (counterbalanced):

1. Consensus Estimation

Out of 100 people, how many people would say this is ***a [ADJ] shirt****?*

2. Faultless Disagreement

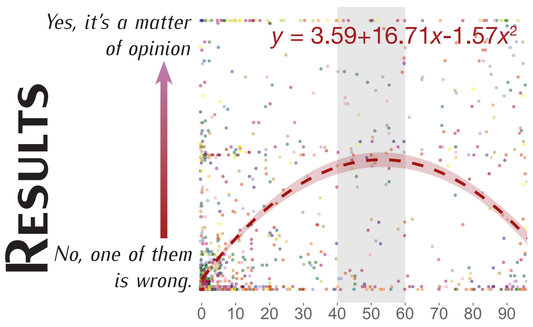
Two people, A and B, are looking at this shirt.

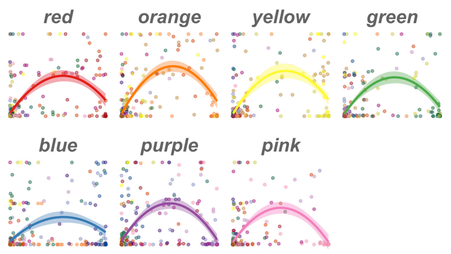
A says, "That's ***a [ADJ] shirt.****"*
B says, "No it's not! That's not ***a [ADJ] shirt****."*

Can both be right?

Results

Color Terms



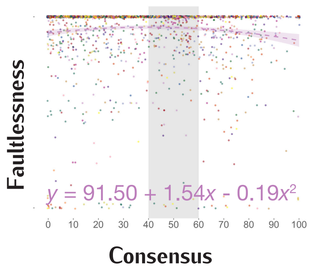


Quadratic relation[5] between estimated consensus and faultless disagreement judgments for both color terms ($X^2(1)$=479.19, p<.001) and evaluative predicates ($X^2(1)$=14.90, p<.001)

Experiment 2 *Are evaluative predicates categorically distinct from color terms (does consensus not matter)?*

Weaker relation between consensus and faultlessness for *e.g., nice, ugly*

Experiment 1



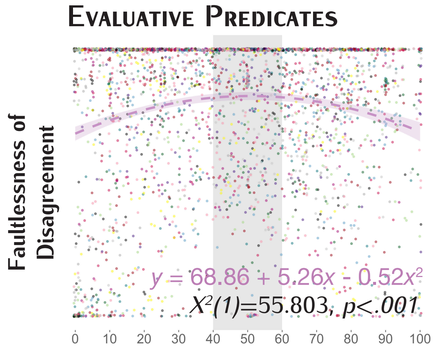
Method

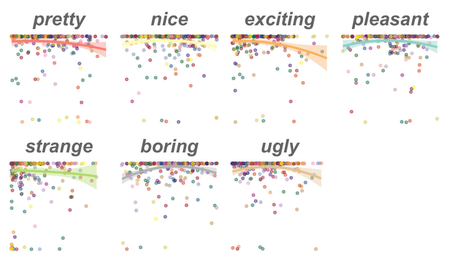
Participants 124 English-speaking Mechanical Turkers

Stimuli 14 'divisive' t-shirt images to elicit greater range of consensus & faultless disagreement judgments for evaluative predicates

Results

Evaluative Predicates





Even when stimuli elicit a range of consensus estimates, speakers tend to judge disagreements over evaluative predicates as 'faultless'

Experiment 3 *Is estimated consensus* **causally** *related to faultless disagreement?*

Method

Participants 160 MTurkers

Stimuli 103 predicate-shirt combinations with ~50% mean consensus estimates in Experiments 1 & 2

To test causality, evidence for population-level consensus for each stimulus is manipulated *within-subjects* across distant trials:

1A. Low Consensus

Out of 100 people, [40-60%, random] ***said that this was a pretty shirt****.*

2A. Faultless Disagreement

...A says, "That's ***a pretty shirt.****"*
B says, "No it's not! That's not ***a pretty shirt****."*

Can both be right?

1B. High Consensus

Out of 100 people, [0-10% / 90-100%, random] ***said that this was a pretty shirt****.*

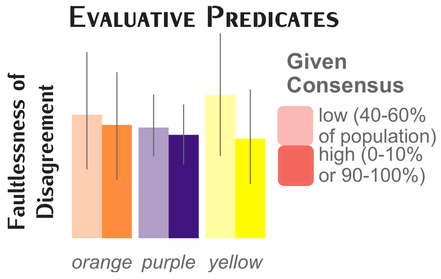
2B. Faultless Disagreement

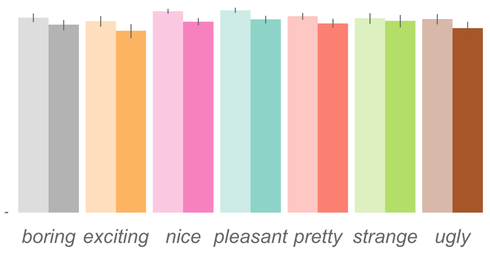
...A says, "That's ***a pretty shirt.****"*
B says, "No it's not! That's not ***a pretty shirt****."*

Can both be right?

Results

Evaluative Predicates





Individual disagreements *over the same stimuli* were less likely to be judged 'faultless' following evidence of high population-level consensus (b=-3.75; $X^2(1)$=46.34, p<.001)

Experiment 4 *Do consensus & subjectivity judgments derive from uncertainty?*

Method

Participants 492 MTurkers

4Blocks (counterbalanced):

1. **Consensus Estimation**
2. **Faultless Disagreement**
3. **Uncertainty Self-Report**
   How certain are you that this shirt is nice?
4. **Explicit Subjectivity**
   Something that is subjective is based on personal opinions or beliefs, rather than objective facts...

   How subjective is the statement that ***this shirt is nice****?*

Results

Consensus and self-reported uncertainty highly correlated (Pearson's r=.65, p<.001)

Nonetheless, estimates of population-level consensus better predict faultless disagreement and subjectivity judgments than epistemic uncertainty.

In Sum

Intuitions about subjectivity and linguistic 'fault' may (sensibly) derive from intuitions about context-specific usage by the speech community

But not the whole story: strength of relation varies across semantic categories...

Future

Can linguistic fault be modeled like moral blame?[7]

Potential continuity between children and adults: Could the locus of children's difficulty with vague or subjective predicates[8] lie in their simulations of the speaker population?

[1] Clark E.V. (1983). https://doi.org/10.1007/978-3-642-69000-6_5
[2] Scontras, G., Degen, J., & Goodman, N. (2017). doi:10.1162/opmi_a_00005
[3] Barker, C. (2013).
[4] Cook, R., Kay, P., & Regier, T. (2003). World Color Survey Naming Data http://www.icsi.berkeley.edu/wcs/data.html
[5] lmer(faultless ~ consensus + consensus2 + (1|shirt) + (1|adjective); Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). 10.18637/jss.v067.i01.
[7] Gerstenberg, T., & Lagnado, D. (2013).
[8] Foushee, R. & Srinivasan, M. (2017).