What Are You Hiding?
Algorithmic Transparency and User Perceptions

Aaron Springer
University of California, Santa Cruz
alspring@ucsc.edu

NETFLIX Browse Kids DVD NEW EPISODES

Because you watched New Girl

TV Mysteries

Irreverent TV Shows

Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

Journal of Economic Perspectives—Volume 31, Number 2—Spring 2017—Pages 211–236

Social Media and Fake News in the 2016 Election

Hunt Allcott and Matthew Gentzkow

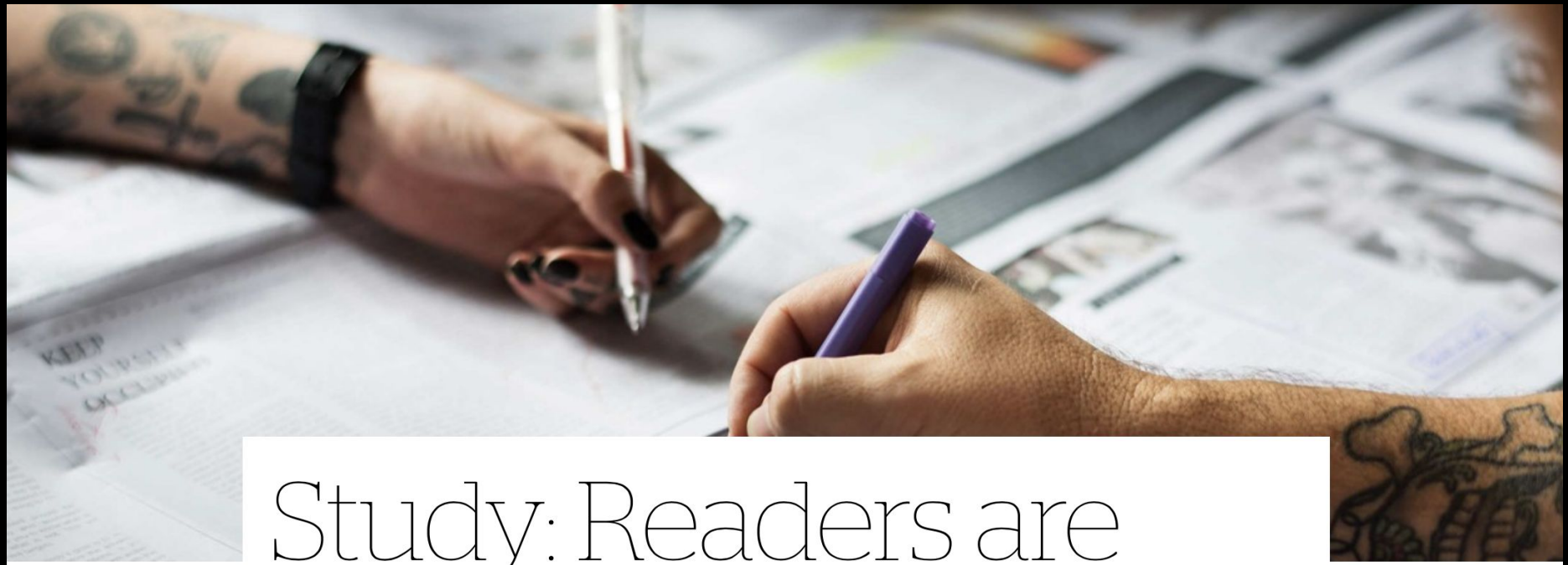
LETTER FROM USACM

Toward Algorithmic Transparency and Accountability

Report calls for algorithmic transparency and education to fight fake news

Natasha Lomas @riptari / Mar 12, 2018

Comment



Study: Readers are hungry for news feed transparency

TOW

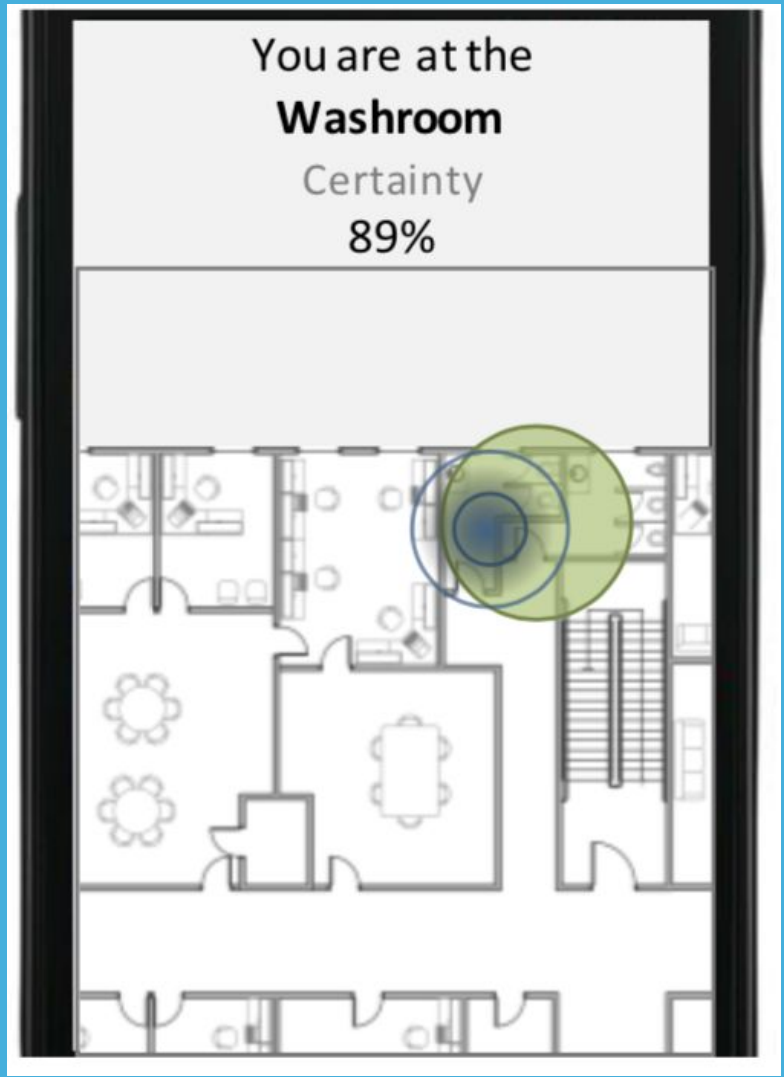
TOW CENTER

How does transparency affect user perceptions of trust and accuracy?

Background

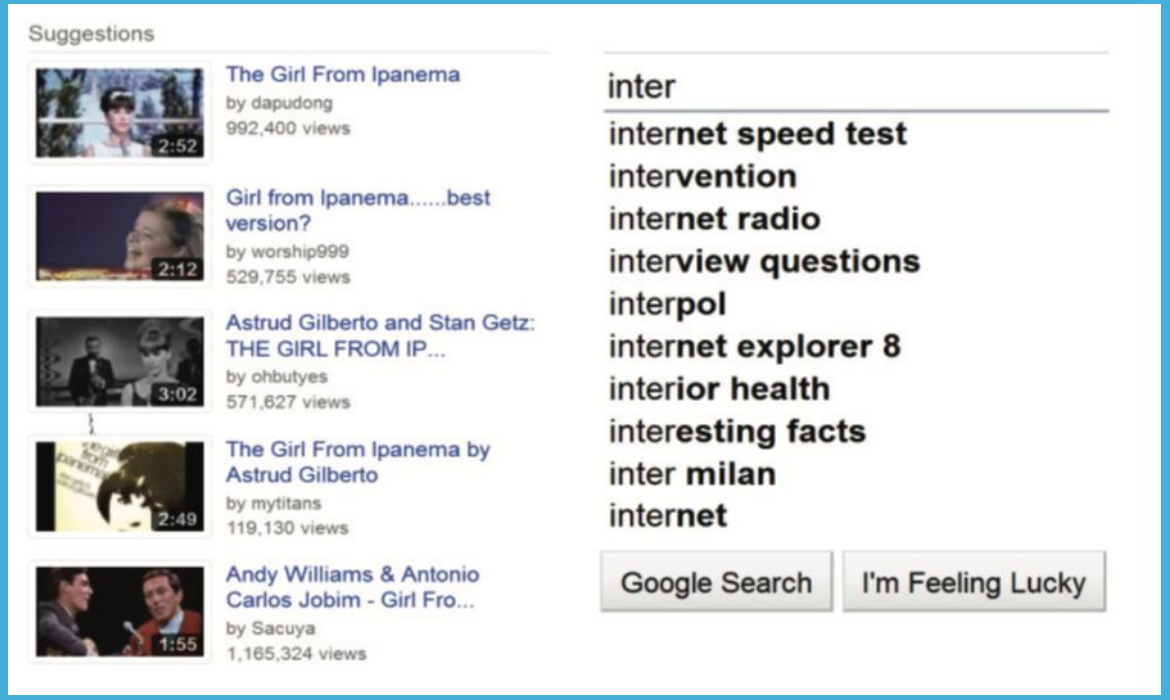
Users may question transparent algorithm when it's correct

(Lim and Dey, 2011)



Or transparency may just create additional processing for users

(Bunt, Lount, Lauzon, 2012)



Transparency may protect trust in the system when user expectations are violated

(Kizilcec, 2016)

EmotiCal (Hollis et al. 2016)

- Perceived accuracy of the algorithms correlated with engagement in application

Experiment

E-meter System

The graphic above displays the output from an algorithm that assesses the positivity/negativity of your writing as you answer the question below. Individual words you type will be highlighted according to whether they contribute towards a positive mood direction or negative mood direction.

Please write at least 100 words about an emotional experience that affected you in the last week.

76/100 words

Earlier today I went hiking with my friend Jonathan. It is always good to spend time with him because he is such a good friend. Though one issue we ran into was that his dog died recently so we were really sad that he had to go through that. Fortunately, I was there for comfort him so things ended up being alright for us—I think he is happy now after we spent that time together.

Experiment Design

- 41 participants from Amazon Turk
- Survey followed consisting of questions assessing accuracy, trust, and corresponding free-writes
- Users were divided into 2 conditions

Control Condition Displays Only the Visualization

The graphic above displays the output from an algorithm that assesses the positivity/negativity of your writing as you answer the question below.

1. For each of the past 3 days: Choose one event that affected you emotionally and write a paragraph about how and why it affected you.

I went to the vet and got some really good news. Baxter is going to be okay after all.

Transparent Condition Shows Contribution of Written Words to Overall Rating

The graphic above displays the output from an algorithm that assesses the positivity/negativity of your writing as you answer the question below. Individual words you type will be highlighted according to whether they contribute towards a positive mood direction or negative mood direction.

Please write at least 100 words about an emotional experience that affected you in the last week.

76/100 words

Earlier today I went hiking with my friend Jonathan. It is always good to spend time with him because he is such a good friend. Though one issue we ran into was that his dog died recently so we were really sad that he had to go through that. Fortunately, I was there for comfort him so things ended up being alright for us—
I think he is happy now after we spent that time together.

Algorithm Decisions are UX Decisions

- Linear Regression
  - Intercept

Algorithm Decisions are UX Decisions

- Features are stemmed unigrams

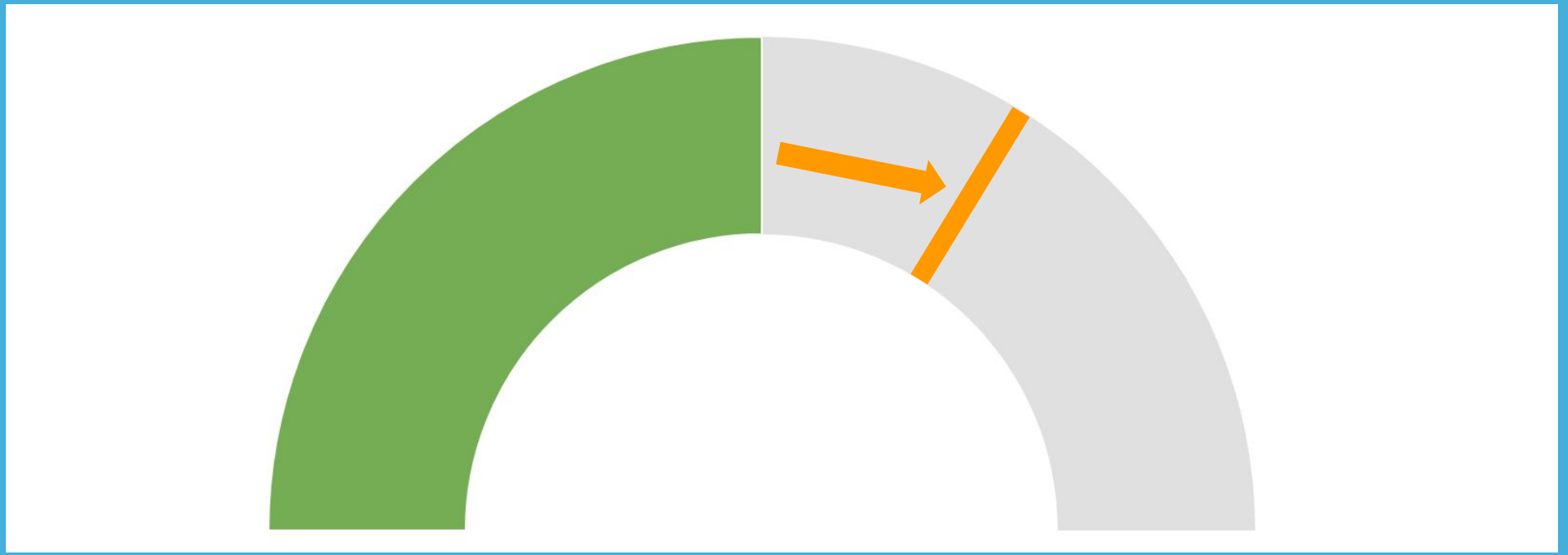
Algorithm Decisions are UX Decisions

- 600 word vocabulary (stemmed)
  - Top features selected by F-score
  - Performance/accuracy trade-off

Algorithm Decisions are UX Decisions

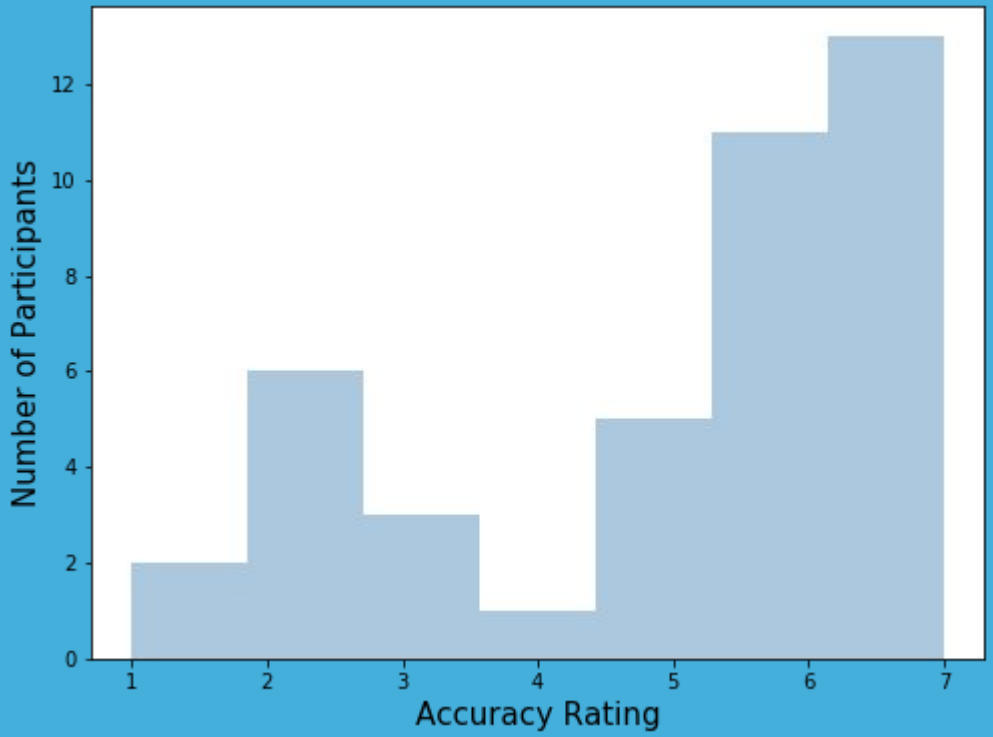
- Mean Absolute Error = 0.94 (on [-3, 3] scale)

Algorithm Average Error



Results

Both conditions found the E-meter accurate



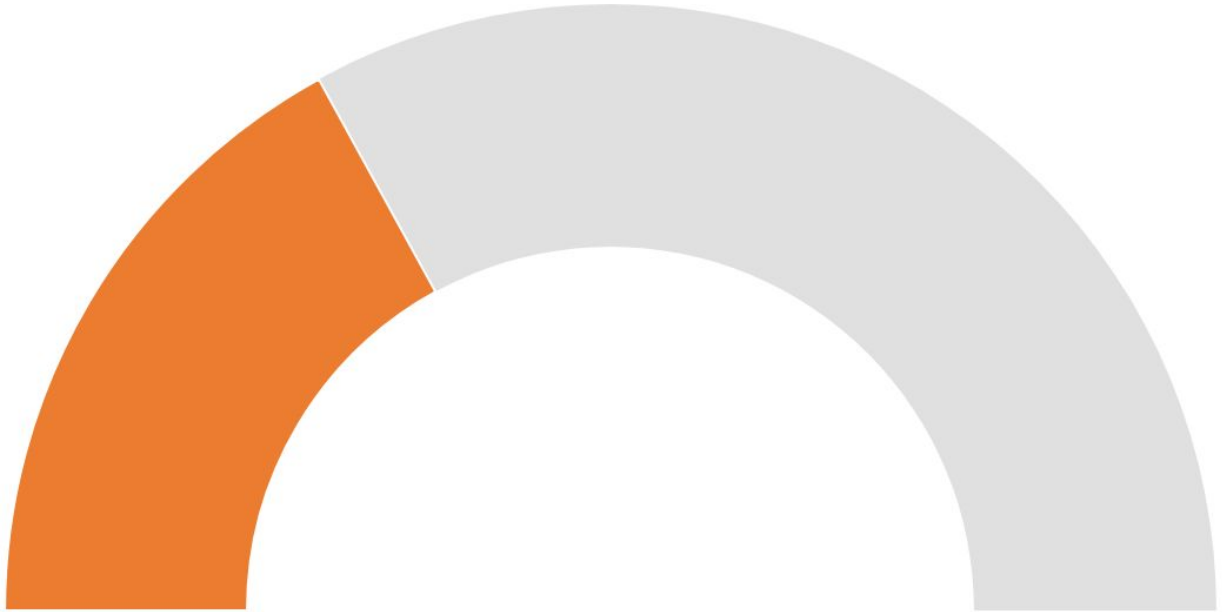
Scale from Very Inaccurate to Very Accurate

Participants were moderately trusting of the E-meter

Scale from Not At All Trustworthy to Very Trustworthy

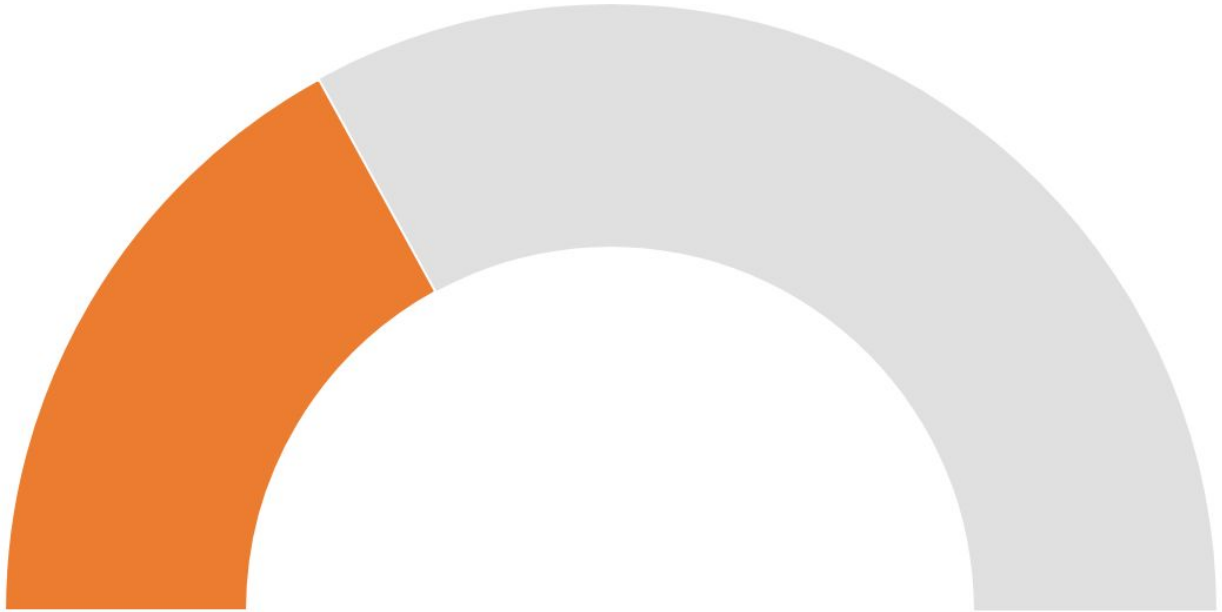
What is expectation violation?

User Expectation

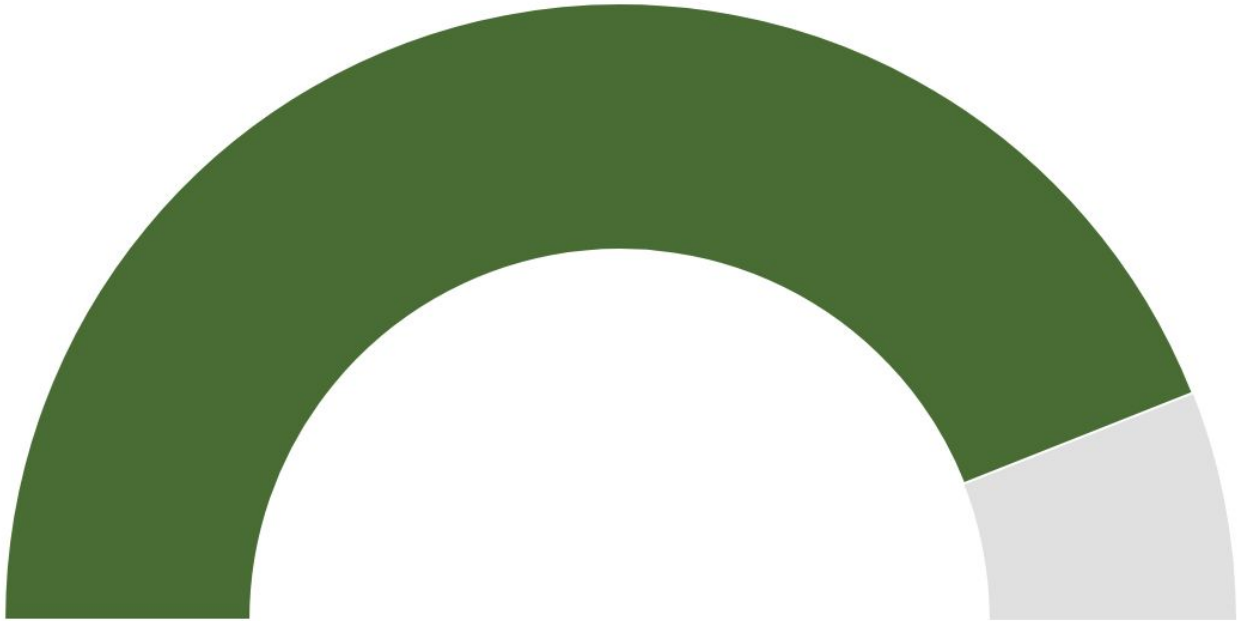


System Output

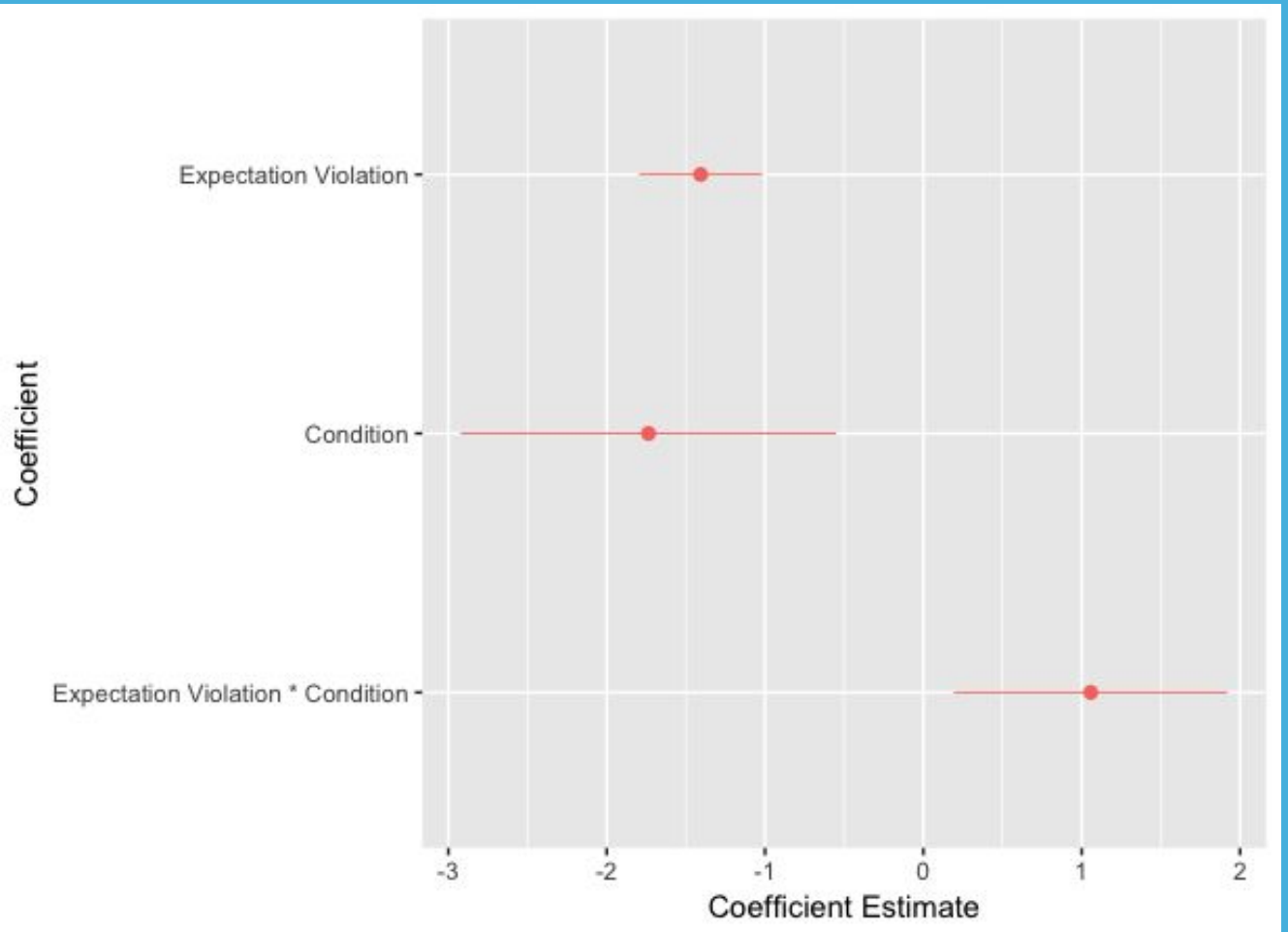
User Expectation



System Output



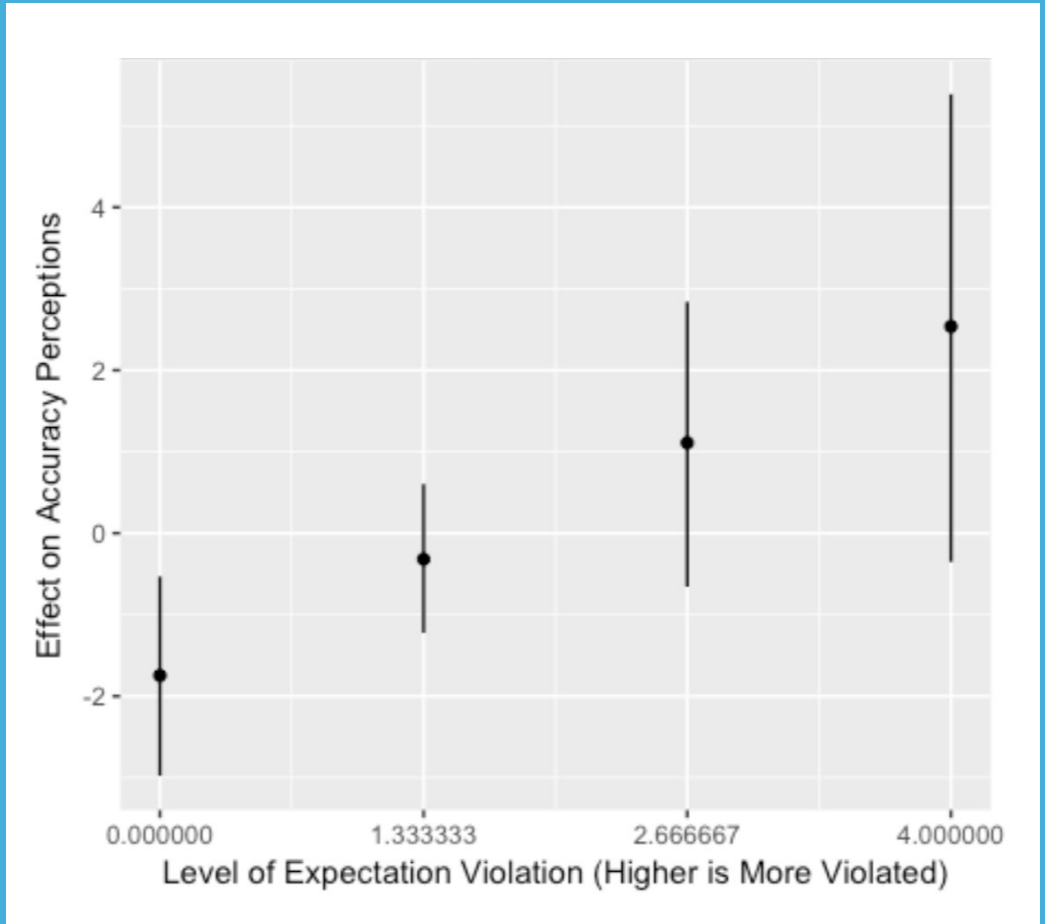
Accuracy Perceptions Regression



$p < .0001$

$R^2 = .548$

Transparency Reduces Accuracy Perceptions Except When Expectations are Violated



Why do users with transparency perceive the E-meter as inaccurate even when it is correct?

Lack of Personalization

“Family is bad for me but it was marked in green”

- Participant 34

Questions Created By Transparency

“The rating does not correspond to the number of identified words”

- Participant 39

Conclusions

- Algorithm Decisions are UX Decisions

Conclusions

- Transparency itself is complex and has a complex relationship with the user experience

Conclusions

- Let’s explore selective transparency

Let’s talk! - alspring@ucsc.edu

- Algorithm Decisions are UX Decisions
- Transparency itself is complex has a complex relationship with the user experience
- Let’s explore selective transparency