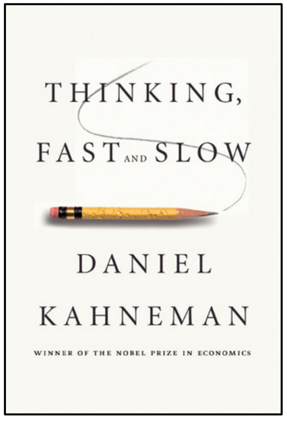
Bayesian reasoning in the social domain

Jack Cao

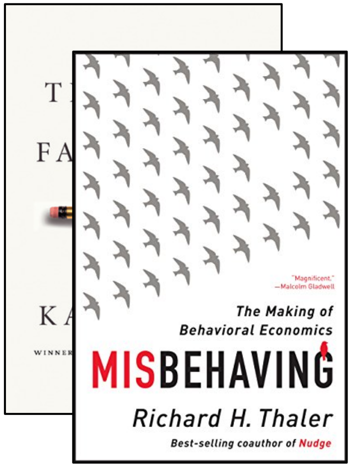
April 24, 2019

Prior Belief x Quality of New Info = Updated Belief

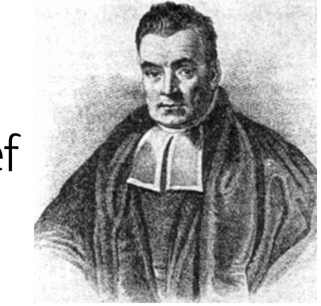
Prior Belief x Quality of New Info = Updated Belief

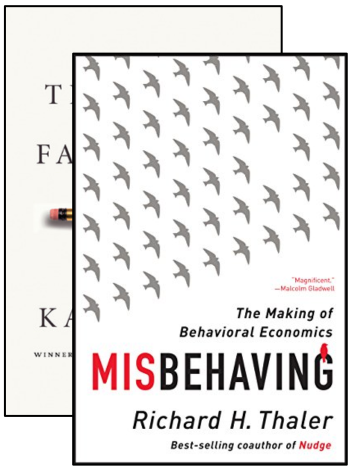


Prior Belief x Quality of New Info = Updated Belief

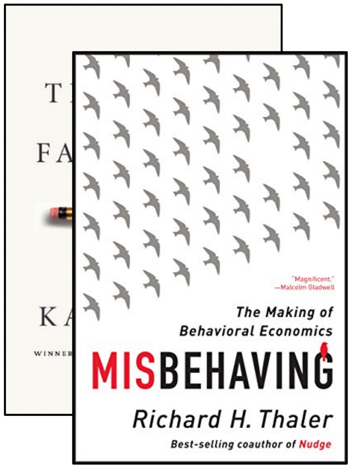


Prior Belief x Quality of New Info = Updated Belief

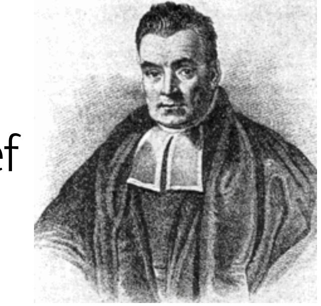


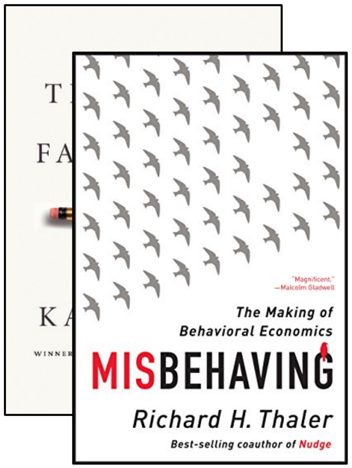


Prior Belief x Quality of New Info = Updated Belief



Prior Belief x Quality of New Info = Updated Belief

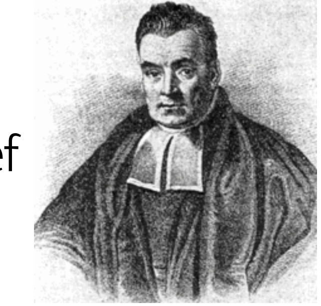




"The theory that would not die..."
Sharon McGrayne, Science Writer

"...arguably the most powerful mechanism created for processing data and knowledge."
Jim Berger, Statistician

Prior Belief x Quality of New Info = Updated Belief



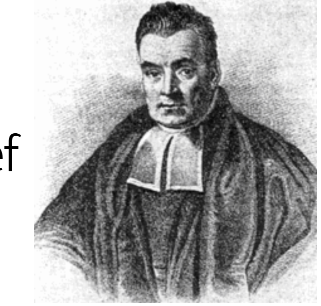
In the social domain, priors are stereotypes.

Locksley et al. (1980)

Krosnick et al. (1990)

Jussim (2012)

Prior Belief x Quality of New Info = Updated Belief

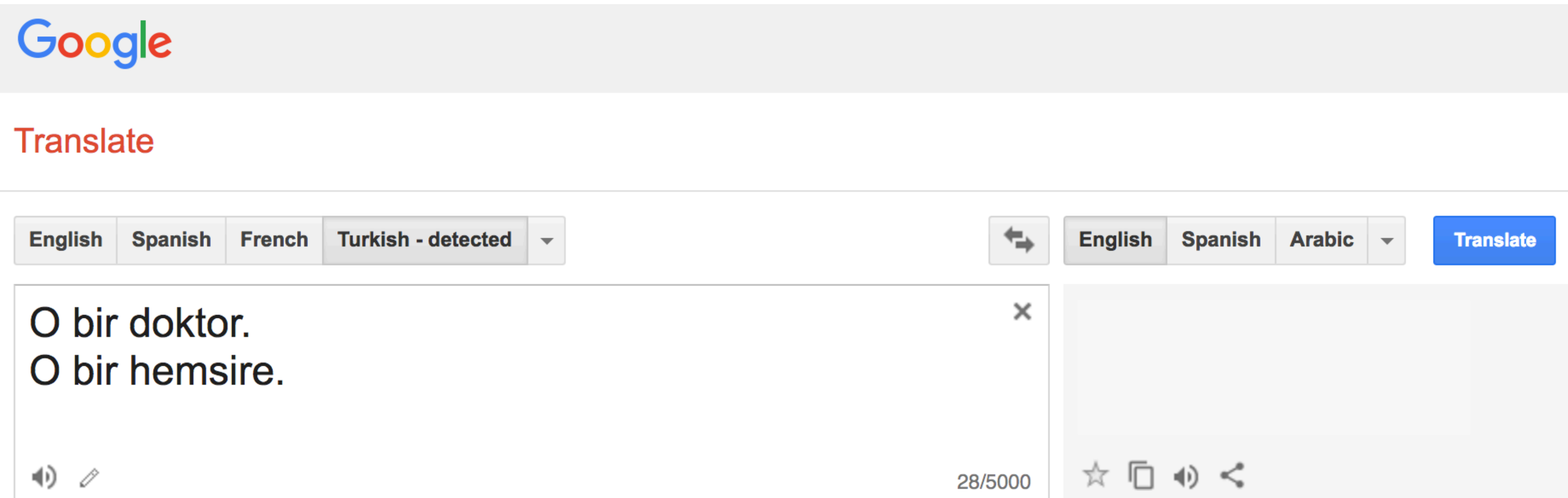


In the social domain, priors are stereotypes.

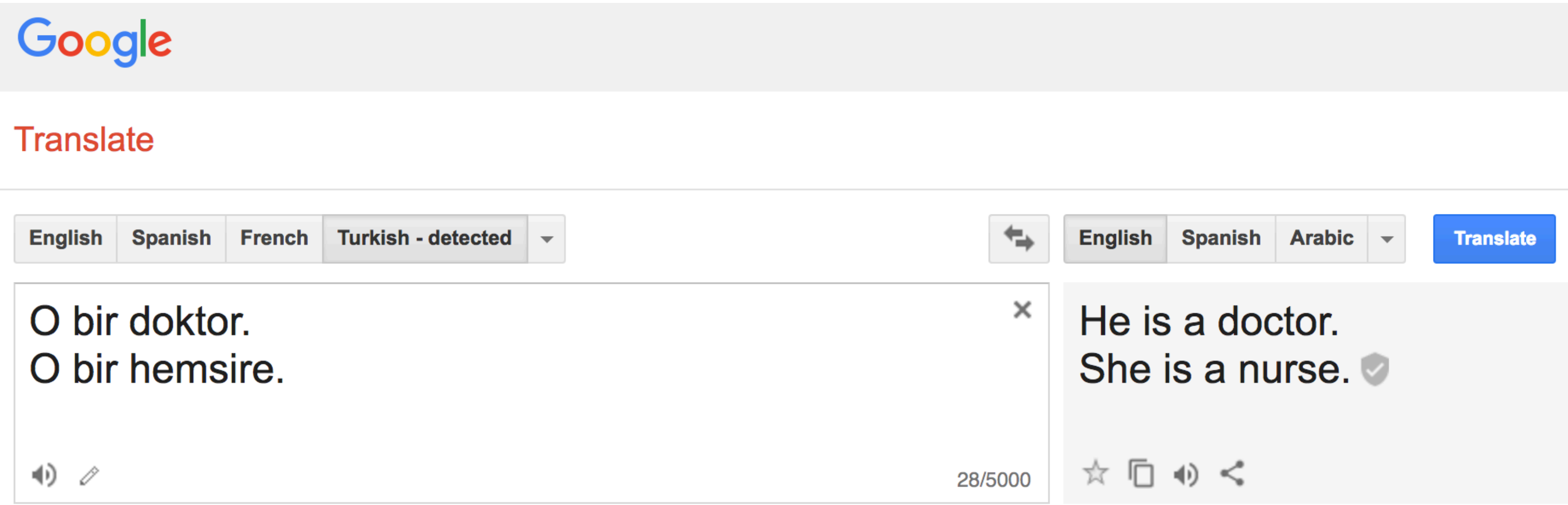
Locksley et al. (1980)
Krosnick et al. (1990)
Jussim (2012)

The law privileges Egalitarian values over Bayesian principles.

Artificial Intelligence
that's more Bayesian than egalitarian



Artificial Intelligence
that's more Bayesian than egalitarian



What judgments do people make when Bayesian principles and egalitarian values are at stake?

What judgments do people make when Bayesian principles and egalitarian values are at stake?

1. People undermine their commitment to egalitarian values by making Bayesian judgments

What judgments do people make when Bayesian principles and egalitarian values are at stake?

1. People undermine their commitment to egalitarian values by making Bayesian judgments

2. Formalizing "egalitarian values"

What judgments do people make when Bayesian principles and egalitarian values are at stake?

1. People undermine their commitment to egalitarian values by making Bayesian judgments

2. Formalizing "egalitarian values"

3. Google Images as a proxy for the social environment

What judgments do people make when Bayesian principles and egalitarian values are at stake?

A man performed surgery.
A woman performed surgery.
Who's more likely to be a doctor?

Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

What judgments do people make when Bayesian principles and egalitarian values are at stake?

A man performed surgery.
A woman performed surgery.
Who's more likely to be a doctor?

Man				
Prior belief	×	Quality of new info	=	Updated belief

Woman				
Prior Belief	×	Quality of new info	=	Updated belief

Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

What judgments do people make when Bayesian principles and egalitarian values are at stake?

A man performed surgery.
A woman performed surgery.
Who's more likely to be a doctor?

Man

Higher × Quality of new info = Updated belief

Woman

Lower × Quality of new info = Updated belief

Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

What judgments do people make when Bayesian principles and egalitarian values are at stake?

A man performed surgery.
A woman performed surgery.
Who's more likely to be a doctor?

Man

Higher × Great, but not perfect = Updated belief

Woman

Lower × Great, but not perfect = Updated belief

Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

Healthcare Network
Nursing in focus

Meet the nurse who will soon perform surgery on patients alone

Unlike other nursing roles, a surgical care practitioner is involved with the patient every step of the way

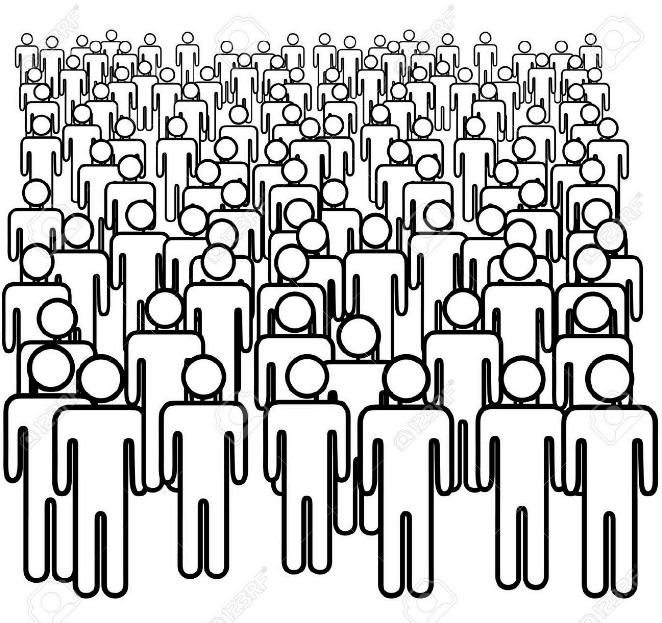
What judgments do people make when Bayesian principles and egalitarian values are at stake?

A man performed surgery.
A woman performed surgery.
Who's more likely to be a doctor?

Man						Woman				
Higher	×	Great, but not perfect	=	More likely		Lower	×	Great, but not perfect	=	Less likely

Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

What judgments do people make when Bayesian principles and egalitarian values are at stake?

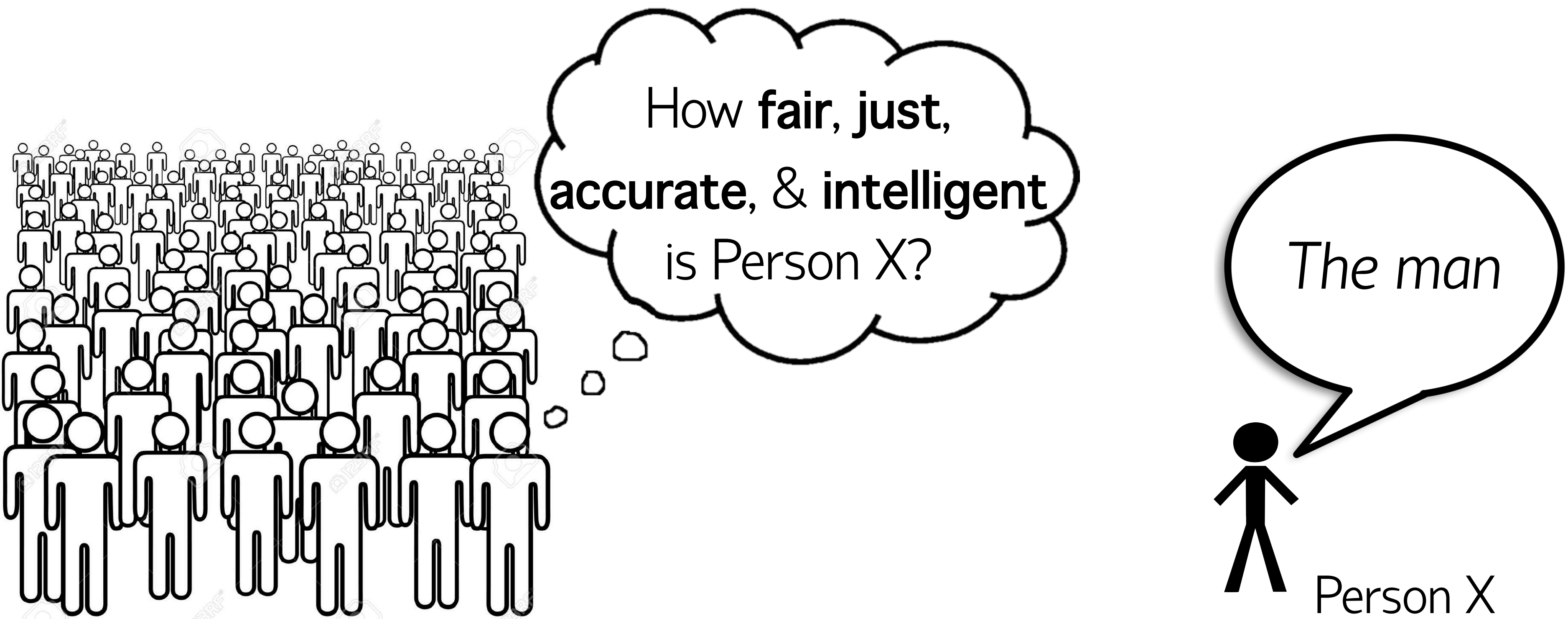


199 MTurkers

Person X

Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

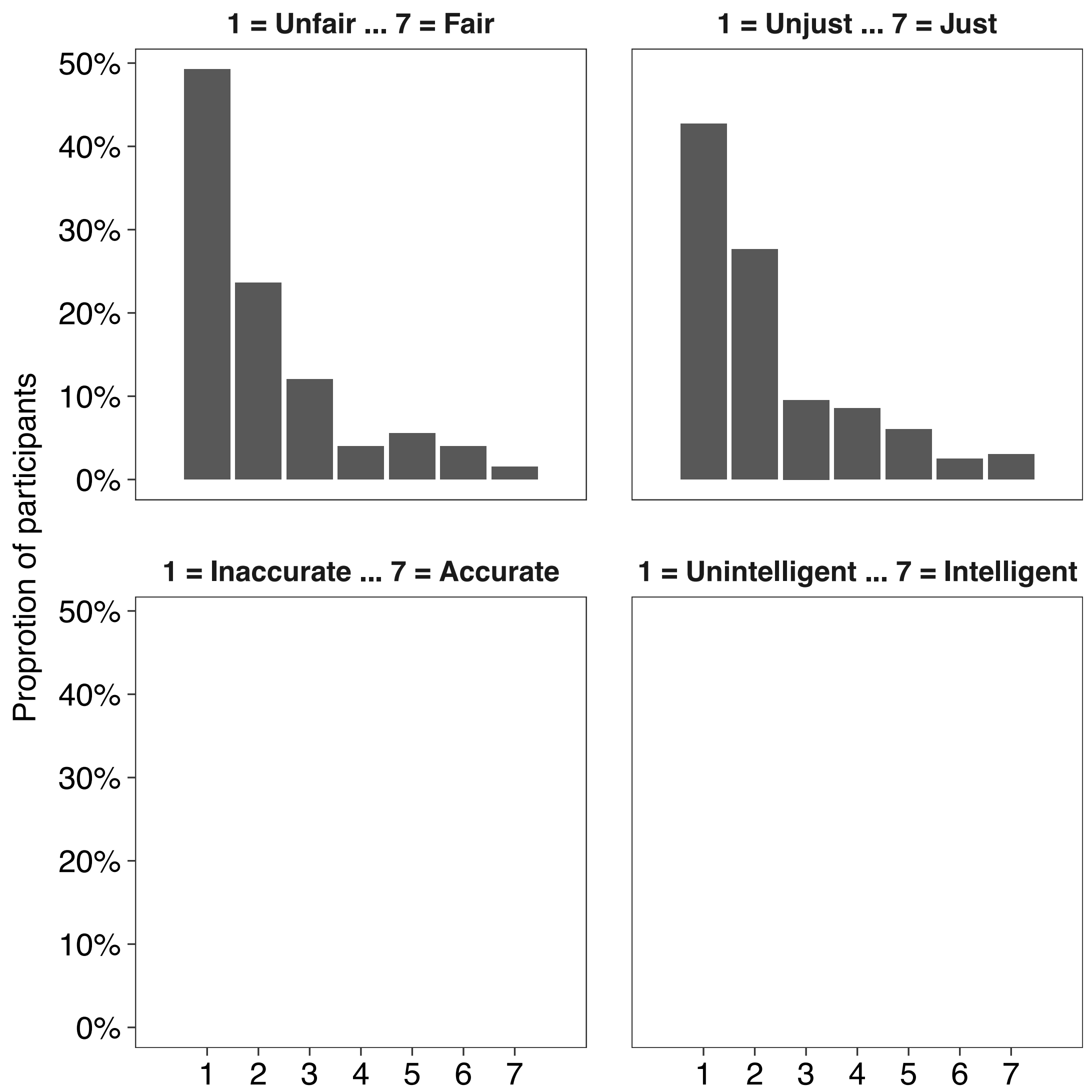
What judgments do people make when Bayesian principles and egalitarian values are at stake?



199 MTurkers

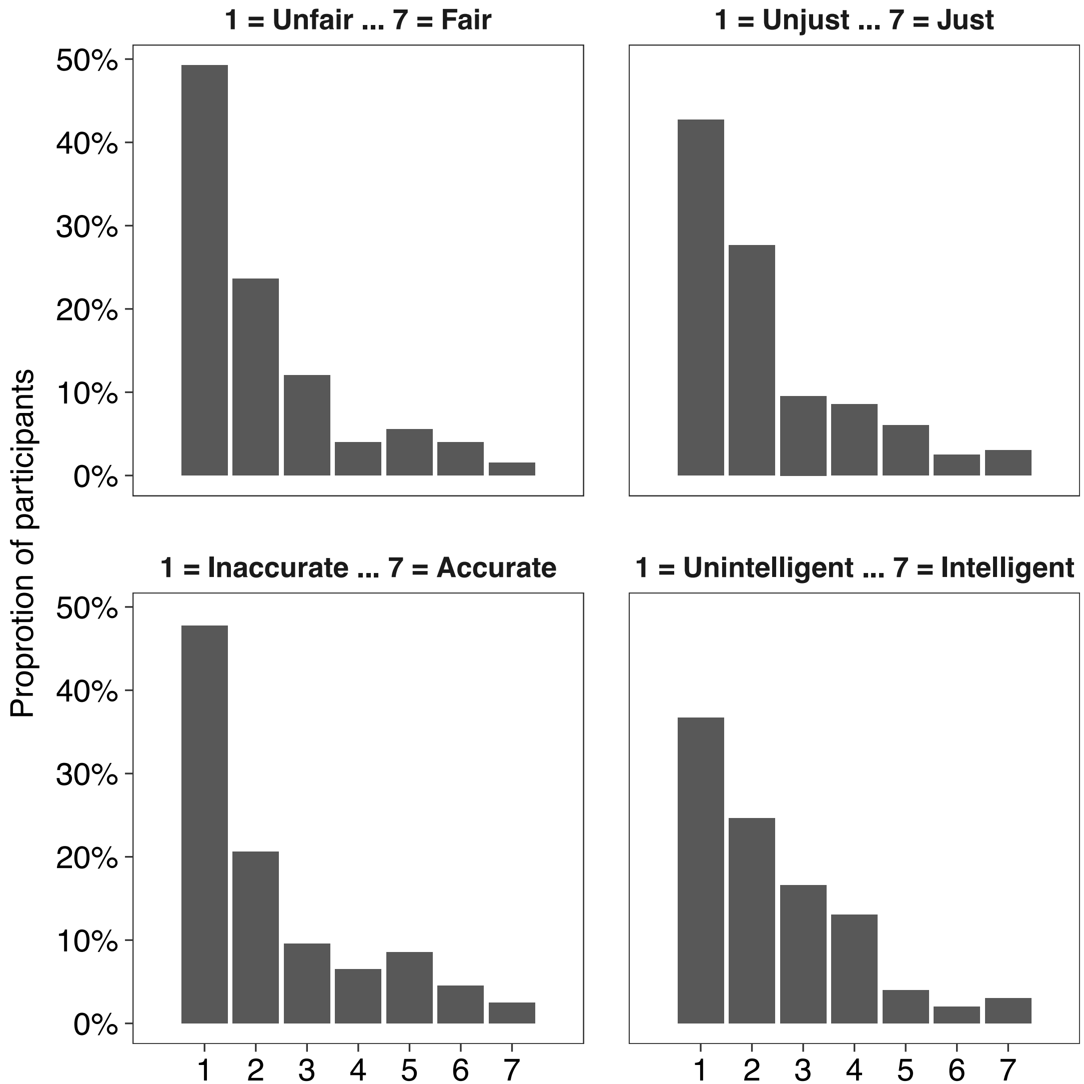
Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

What judgments do people make when Bayesian principles and egalitarian values are at stake?

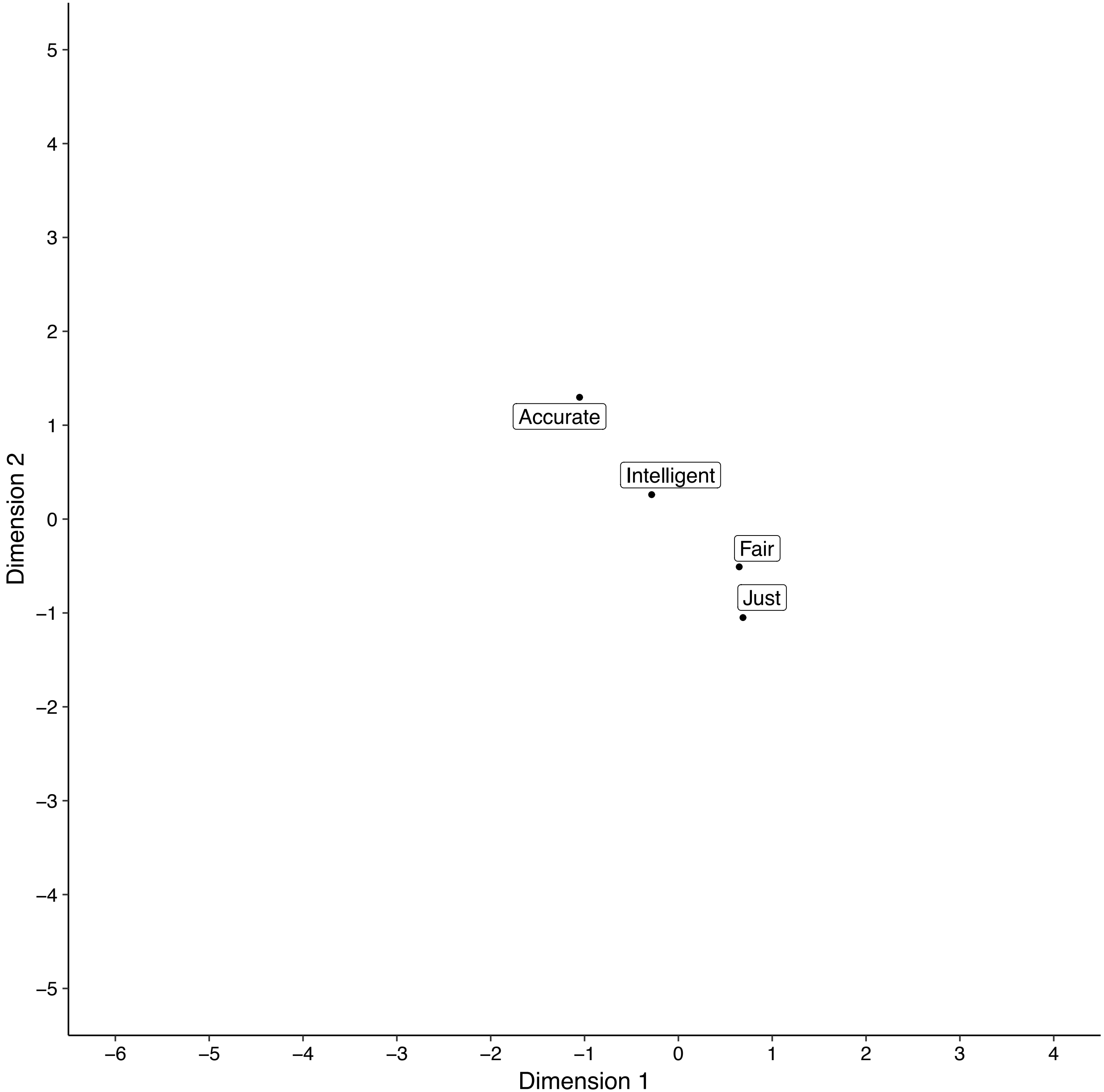


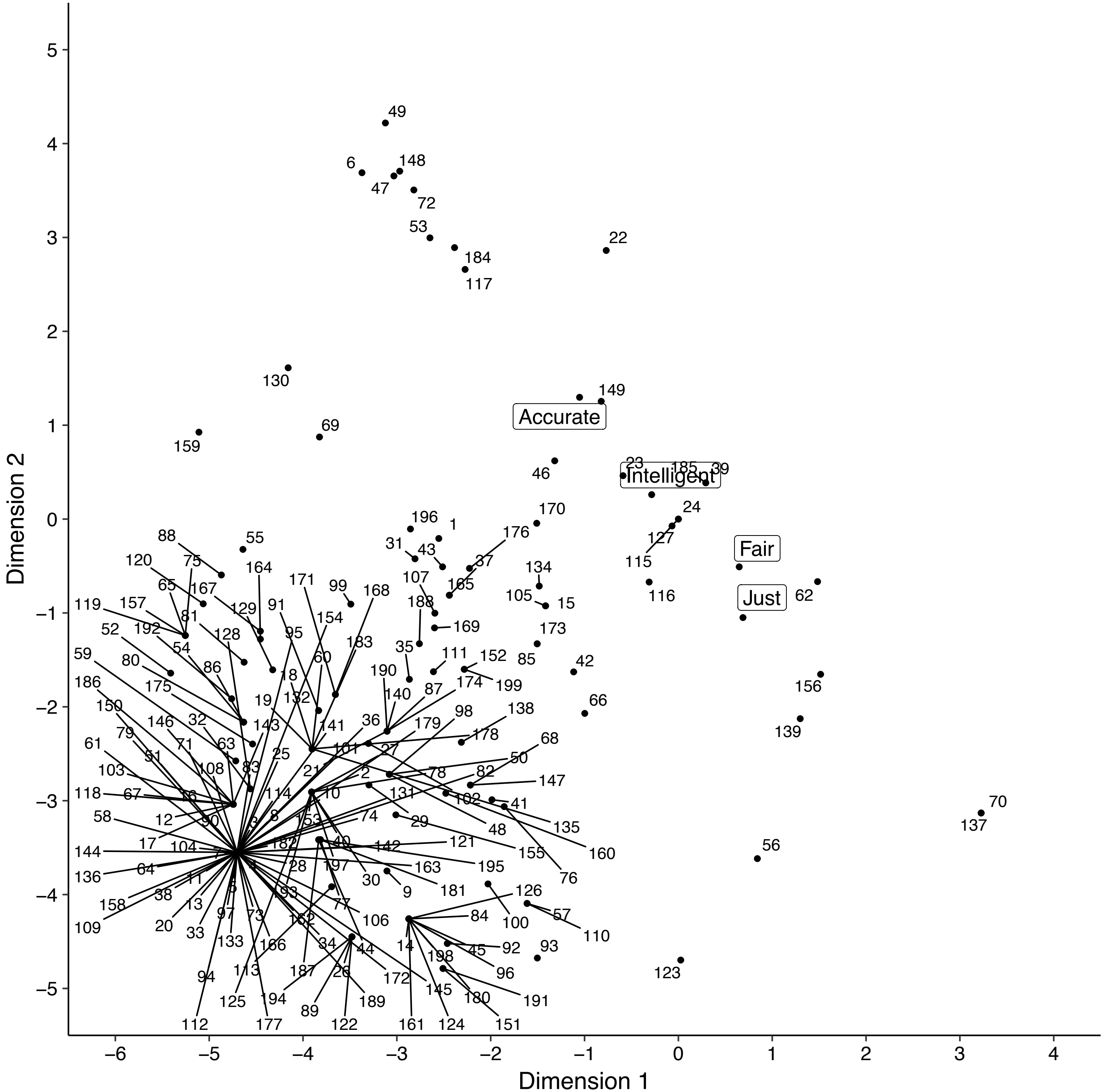
Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

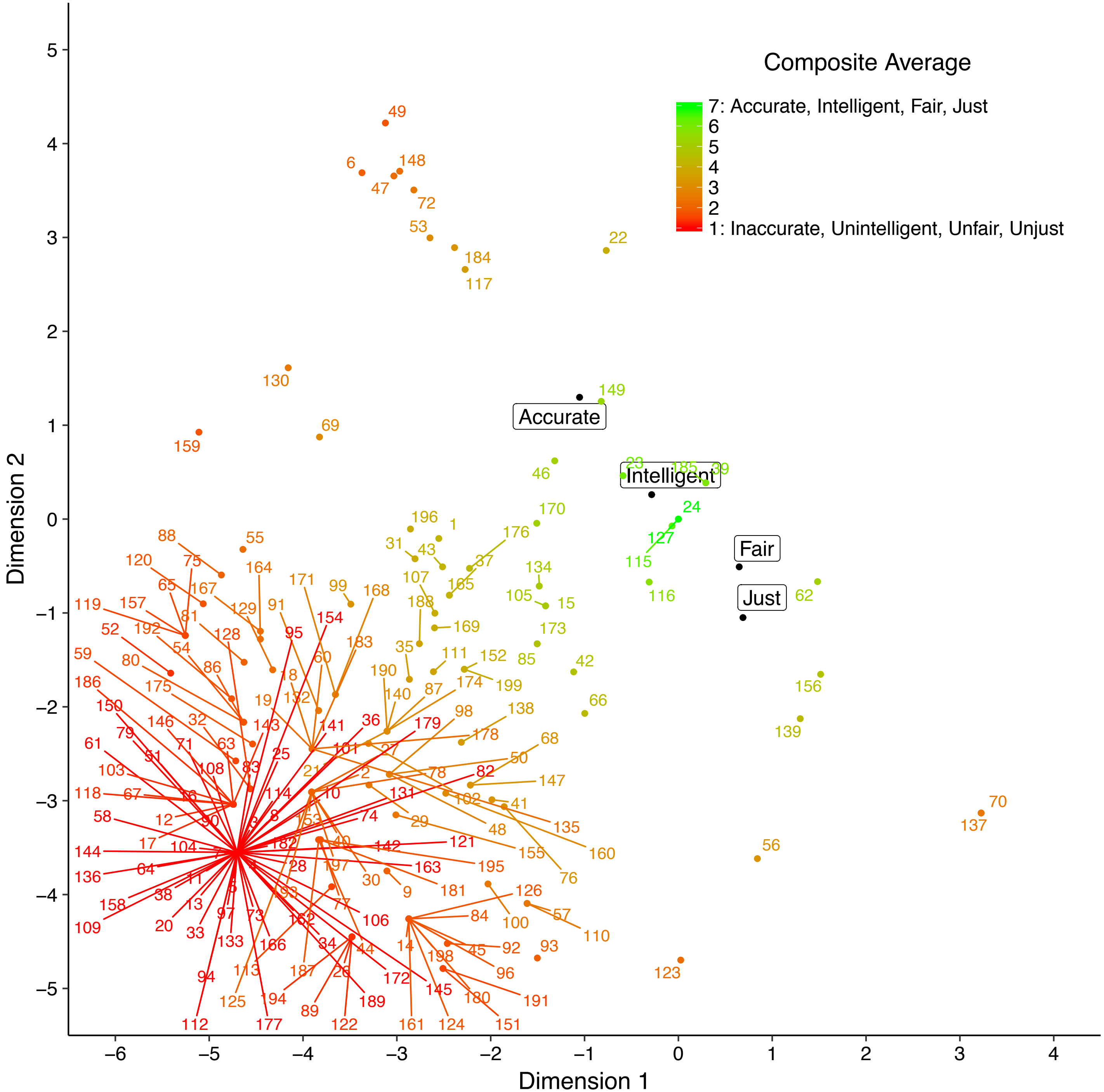
What judgments do people make when Bayesian principles and egalitarian values are at stake?



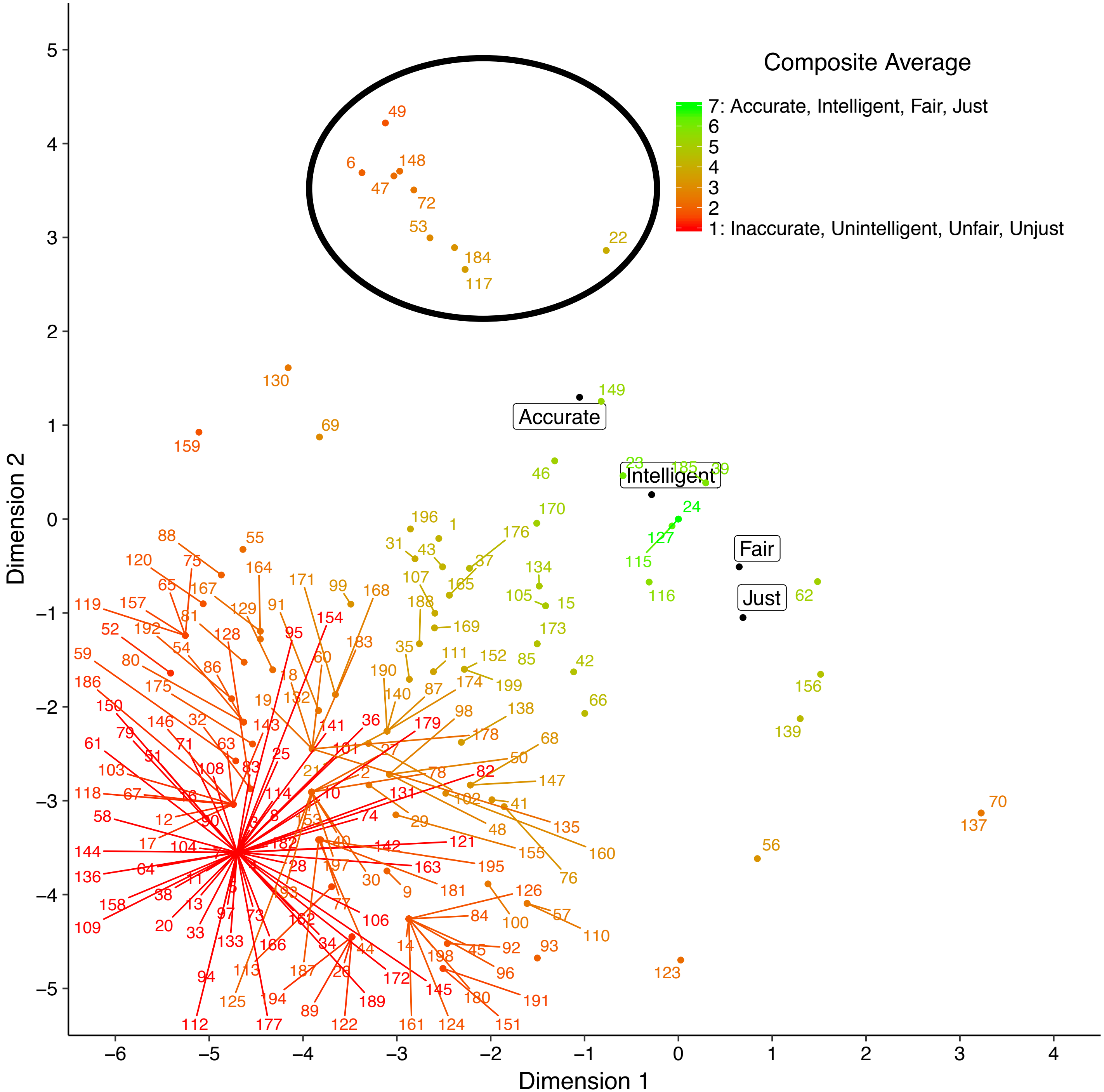
Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

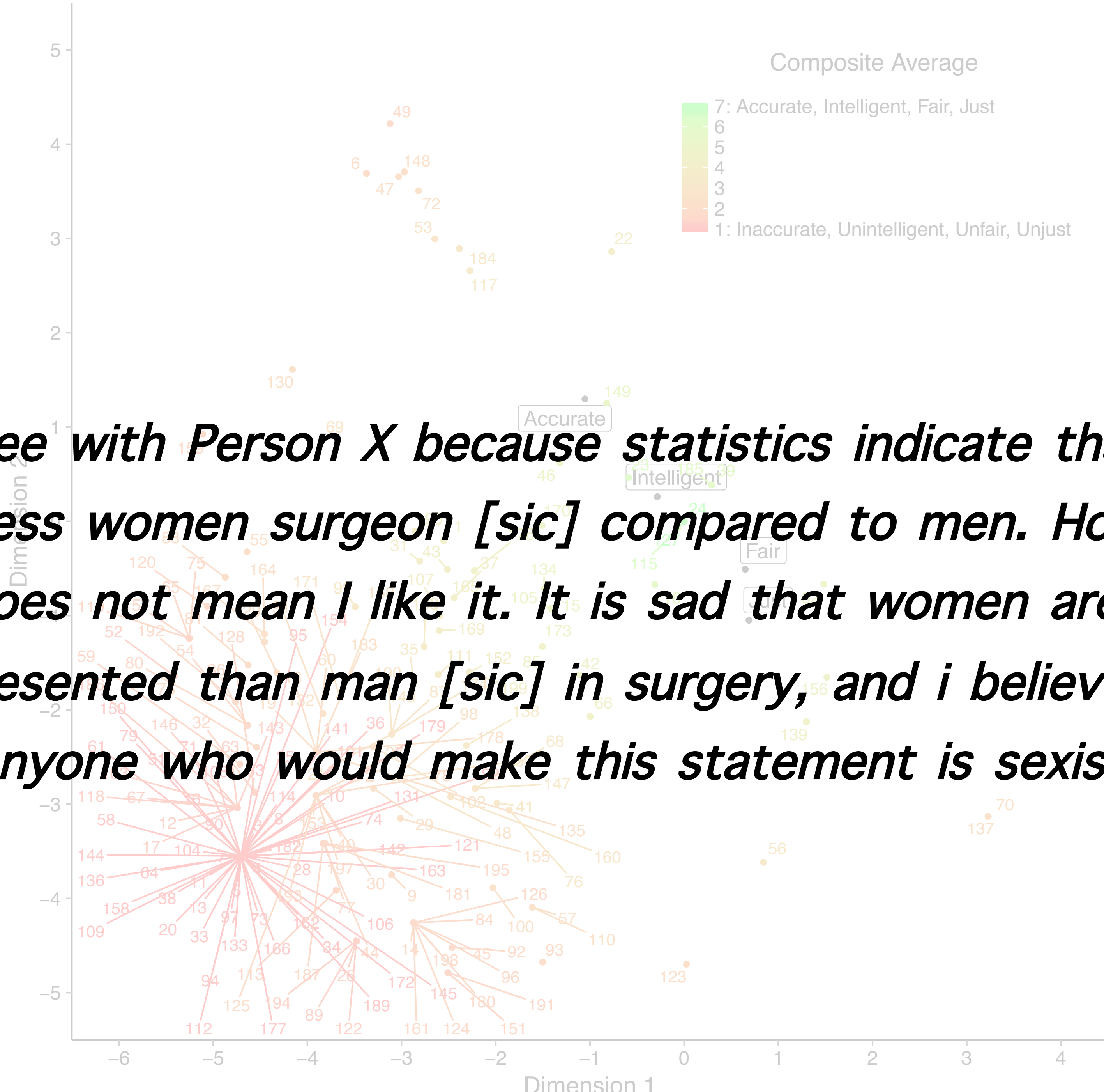

Accurate
Intelligent
Fair
Just
Dimension 1
Dimension 2
−6
−5
−4
−3
−2
−1
0
1
2
3
4
5


Dimension 2
Dimension 1
Accurate
Intelligent
Fair
Just

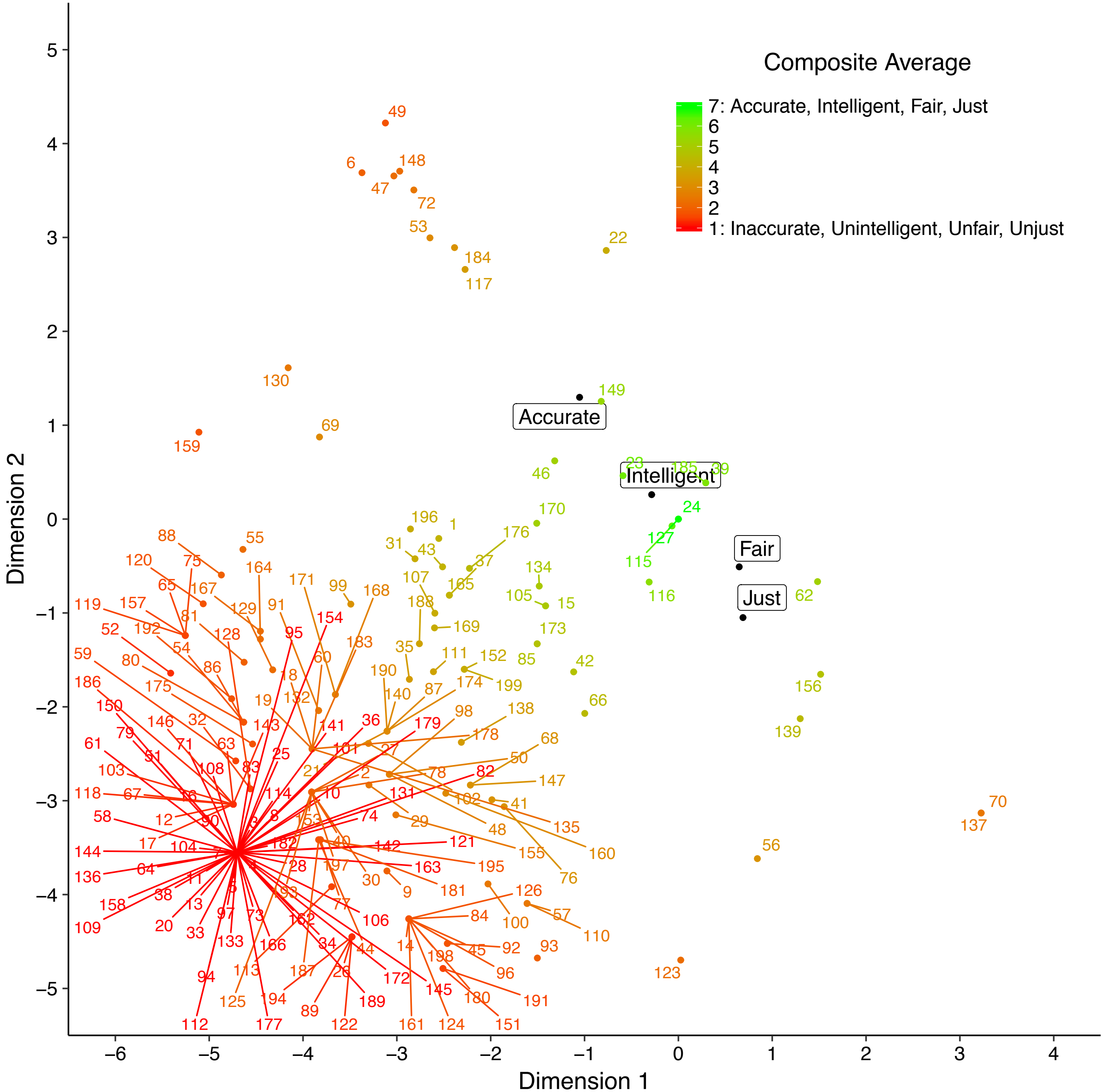


Composite Average
7: Accurate, Intelligent, Fair, Just
6
5
4
3
2
1: Inaccurate, Unintelligent, Unfair, Unjust
Accurate
Intelligent
Fair
Just
Dimension 1
Dimension 2

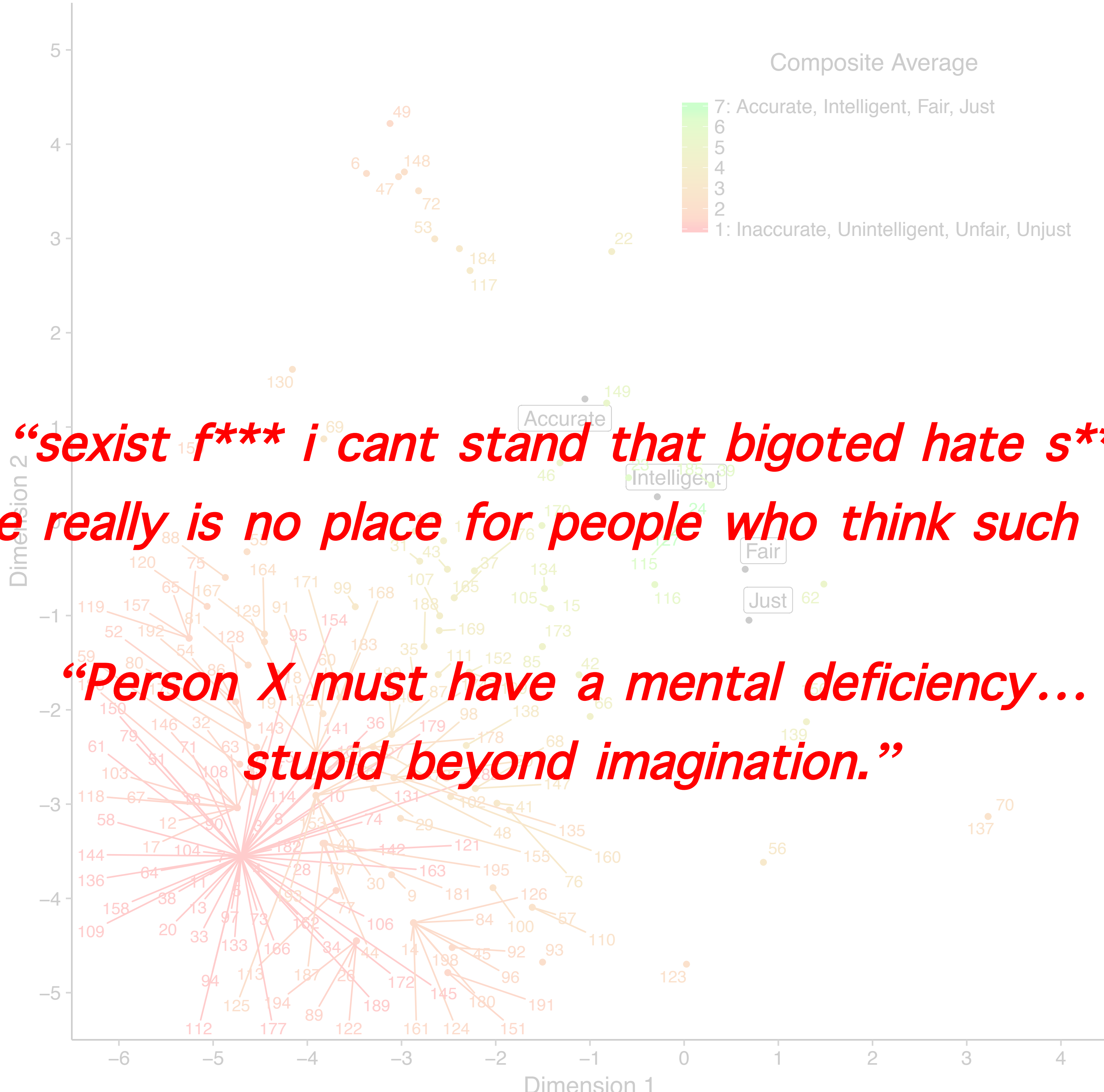

Composite Average
7: Accurate, Intelligent, Fair, Just
6
5
4
3
2
1: Inaccurate, Unintelligent, Unfair, Unjust
Accurate
Intelligent
Fair
Just
Dimension 1
Dimension 2



“I agree with Person X because statistics indicate that there are less women surgeon [sic] compared to men. However, this does not mean I like it. It is sad that women are under-represented than man [sic] in surgery, and i believe that anyone who would make this statement is sexist.”

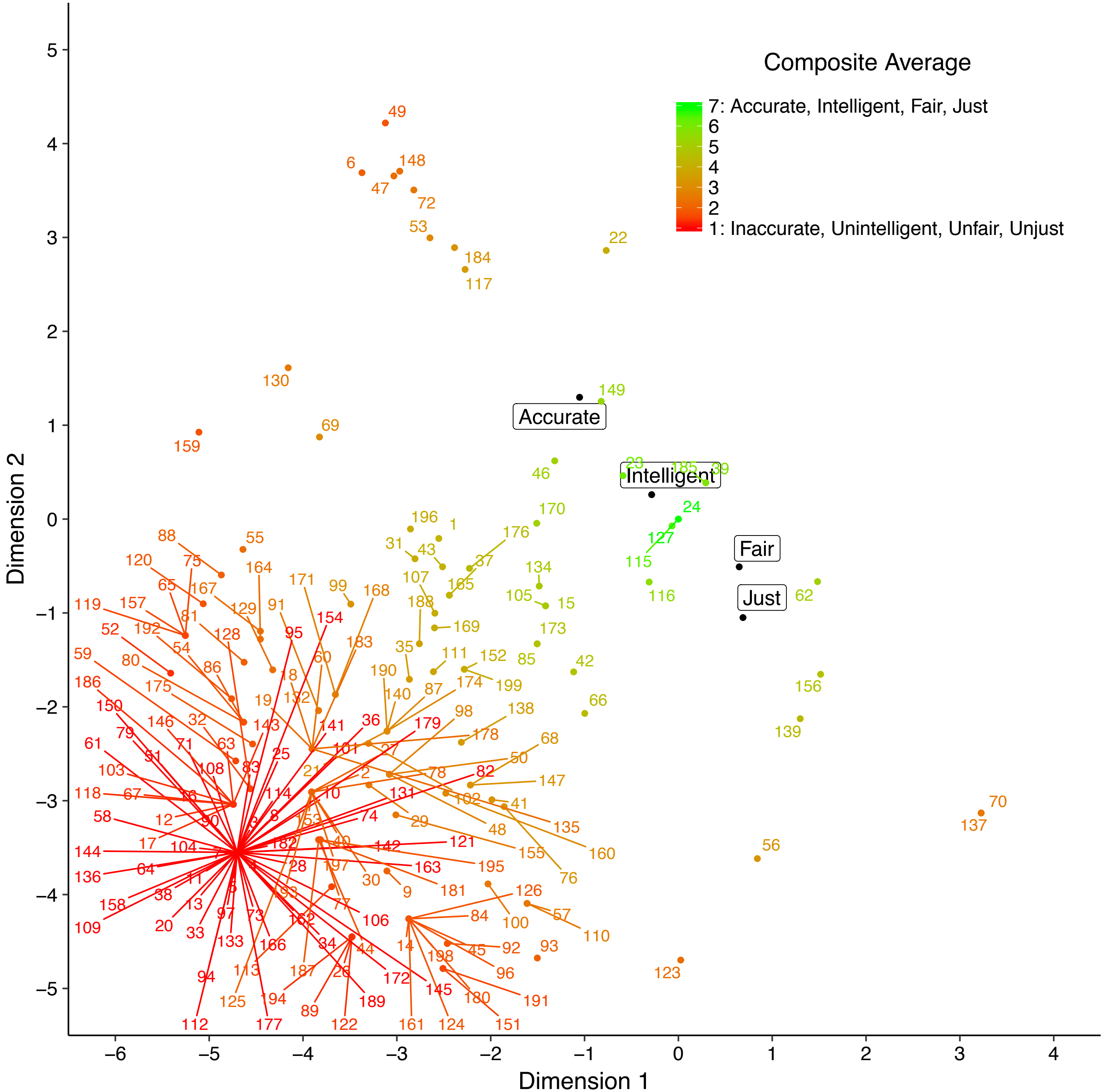


Composite Average
7: Accurate, Intelligent, Fair, Just
6
5
4
3
2
1: Inaccurate, Unintelligent, Unfair, Unjust
Accurate
Intelligent
Fair
Just
Dimension 1
Dimension 2
−6
−5
−4
−3
−2
−1
0
1
2
3
4
5



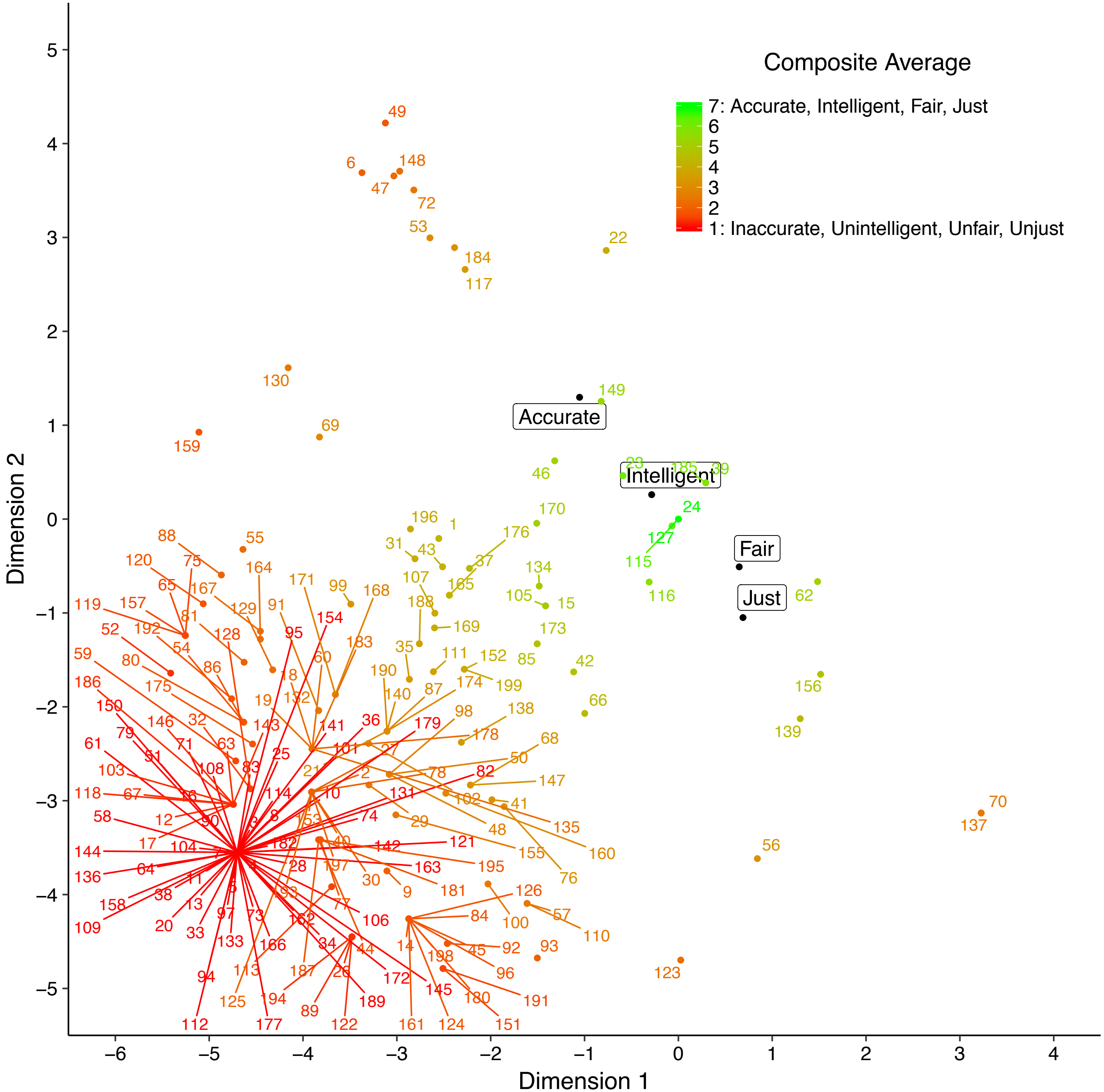
“sexist f i cant stand that bigoted hate s*** there really is no place for people who think such ways”***

“Person X must have a mental deficiency… stupid beyond imagination.”



Composite Average
7: Accurate, Intelligent, Fair, Just
6
5
4
3
2
1: Inaccurate, Unintelligent, Unfair, Unjust
Accurate
Intelligent
Fair
Just
Dimension 1
Dimension 2
−6
−5
−4
−3
−2
−1
0
1
2
3
4
5

“person x is a sexist or a statistician”


Composite Average
7: Accurate, Intelligent, Fair, Just
6
5
4
3
2
1: Inaccurate, Unintelligent, Unfair, Unjust
Accurate
Intelligent
Fair
Just
Dimension 1
Dimension 2

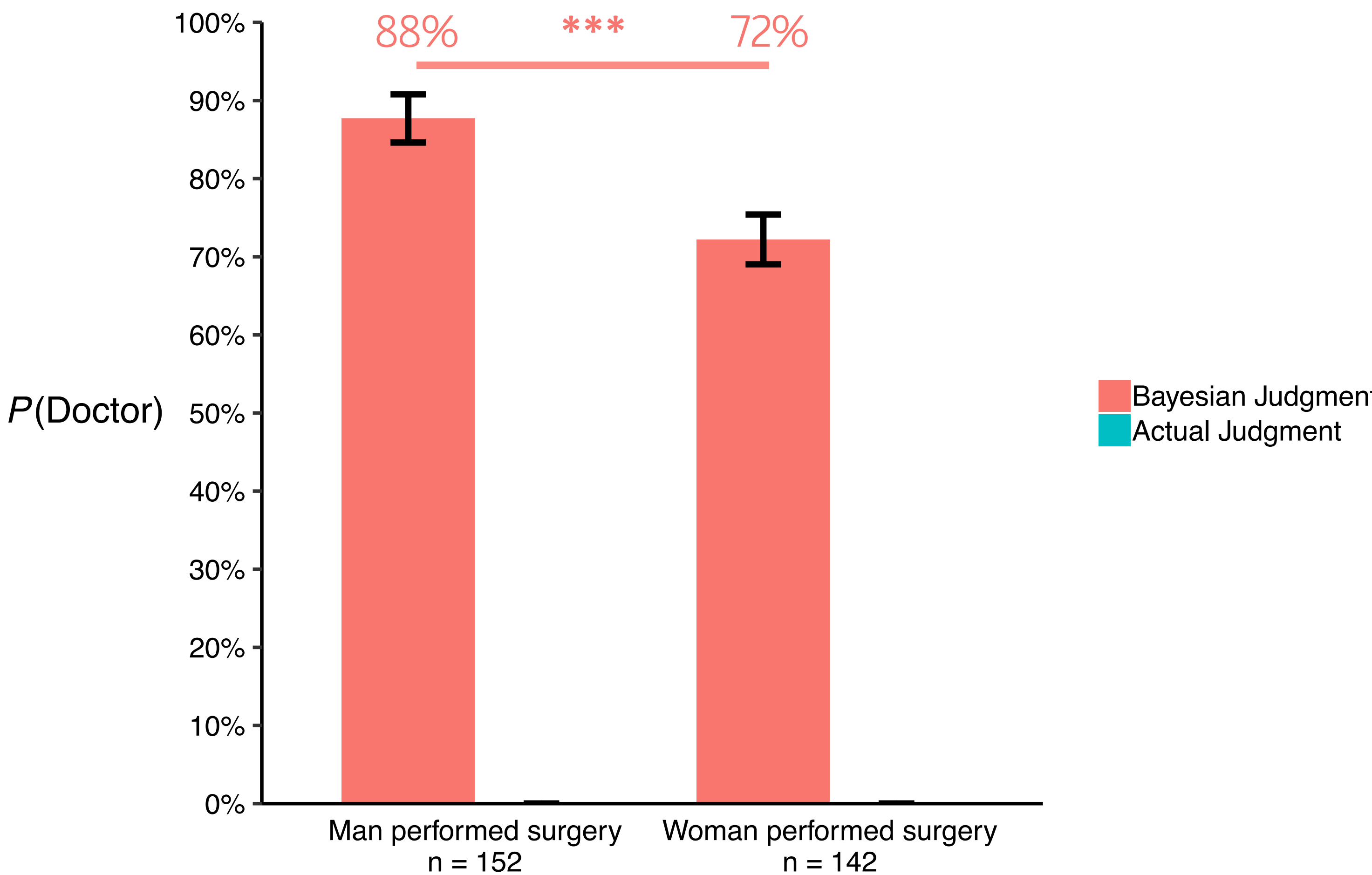
"I just think they're probably not too fun at parties. They're right though, statistically speaking."

What judgments do people make when Bayesian principles and egalitarian values are at stake?

How likely is a man vs. a woman to be a doctor given that each performed surgery?

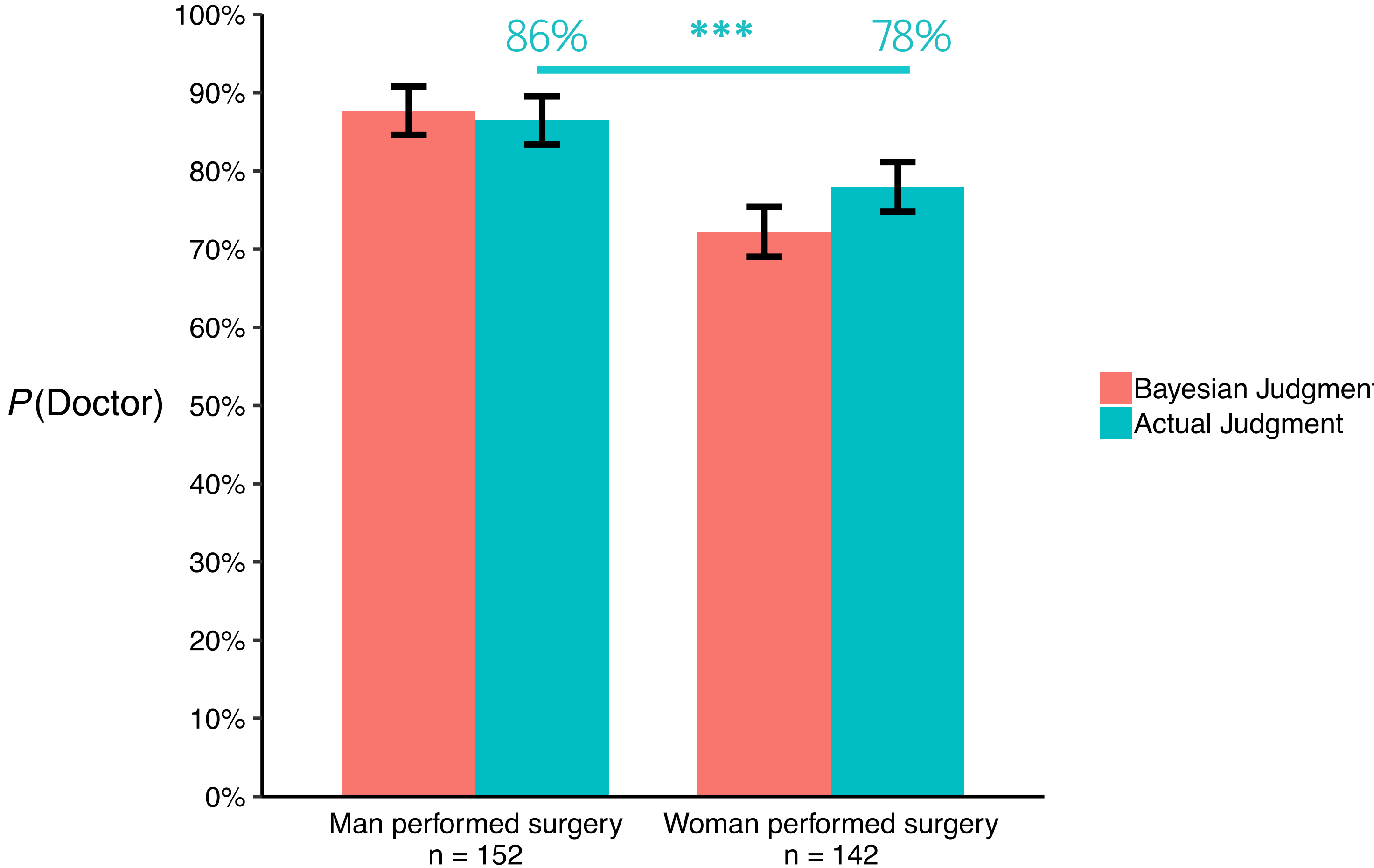
Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

What judgments do people make when Bayesian principles and egalitarian values are at stake?



Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

What judgments do people make when Bayesian principles and egalitarian values are at stake?



Cao, Kleiman-Weiner, & Banaji (2018), *Psych Sci*

What judgments do people make when Bayesian principles and egalitarian values are at stake?

1. People undermine their commitment to egalitarian values by making Bayesian judgments

2. Formalizing "egalitarian values"

3. Google Images as a proxy for the social environment

What judgments do people make when Bayesian principles and egalitarian values are at stake?

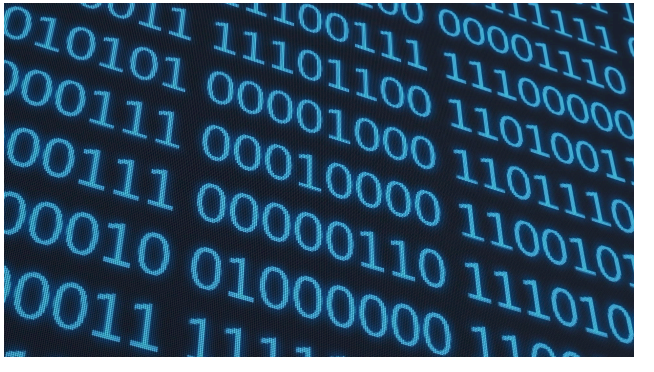
1. People undermine their commitment to egalitarian values by making Bayesian judgments

2. Formalizing "egalitarian values"

3. Google Images as a proxy for the social environment

How do you go from something like this...

to something like this...



What is *un*fair in the social domain?

What is *un*fair in the social domain?

When outcomes depend on social group, *ceteris paribus*

What is *un*fair in the social domain?

When outcomes depend on social group, *ceteris paribus*

- Policing Correll et al. (2007), *JPSP*
- Medicine Hoffman, Trawalter, et al. (2016), *PNAS*
- Hiring Moss-Racusin et al. (2012), *PNAS*

What is *un*fair in the social domain?

When outcomes depend on social group, *ceteris paribus*

Policing Correll et al. (2007), *JPSP*

Medicine Hoffman, Trawalter, et al. (2016), *PNAS*

Hiring Moss-Racusin et al. (2012), *PNAS*

$P(\text{Shoot} \mid \text{Unarmed, White}) < P(\text{Shoot} \mid \text{Unarmed, Black})$

What is *un*fair in the social domain?

When outcomes depend on social group, *ceteris paribus*

Policing Correll et al. (2007), *JPSP*

Medicine Hoffman, Trawalter, et al. (2016), *PNAS*

Hiring Moss-Racusin et al. (2012), *PNAS*

P(Undertreat | In Pain, White) < P(Undertreat | In Pain, Black)

What is *un*fair in the social domain?

When outcomes depend on social group, *ceteris paribus*

Policing Correll et al. (2007), *JPSP*

Medicine Hoffman, Trawalter, et al. (2016), *PNAS*

Hiring Moss-Racusin et al. (2012), *PNAS*

$$P(\text{Hire} \mid \text{Qualified, Female}) < P(\text{Hire} \mid \text{Qualified, Male})$$

$$P(\text{Shoot} \mid \text{Unarmed, White}) < P(\text{Shoot} \mid \text{Unarmed, Black})$$

$$P(\text{Undertreat} \mid \text{In Pain, White}) < P(\text{Undertreat} \mid \text{In Pain, Black})$$

$$P(\text{Hire} \mid \text{Qualified, Female}) < P(\text{Hire} \mid \text{Qualified, Male})$$

Unequal false positive rates, FPRs

P(Shoot | Unarmed, White) < P(Shoot | Unarmed, Black)

P(Undertreat | In Pain, White) < P(Undertreat | In Pain, Black)

P(Hire | Qualified, Female) < P(Hire | Qualified, Male)

<u>Unequal false positive rates, FPRs</u>

P(Shoot | Unarmed, White) < P(Shoot | Unarmed, Black)

<u>Unequal false negative rates, FNRs</u>

P(Undertreat | In Pain, White) < P(Undertreat | In Pain, Black)

P(Hire | Qualified, Female) < P(Hire | Qualified, Male)

Unequal false positive rates, FPRs

P(Shoot | Unarmed, White) < P(Shoot | Unarmed, Black)

Unequal false negative rates, FNRs

P(Undertreat | In Pain, White) < P(Undertreat | In Pain, Black)

Unequal positive predictive values, PPVs

P(Hire | Qualified, Female) < P(Hire | Qualified, Male)

Not egalitarian

<u>Unequal false positive rates, FPRs</u>

$$P(\text{Shoot} \mid \text{Unarmed, White}) < P(\text{Shoot} \mid \text{Unarmed, Black})$$

<u>Unequal false negative rates, FNRs</u>

$$P(\text{Undertreat} \mid \text{In Pain, White}) < P(\text{Undertreat} \mid \text{In Pain, Black})$$

<u>Unequal positive predictive values, PPVs</u>

$$P(\text{Hire} \mid \text{Qualified, Female}) < P(\text{Hire} \mid \text{Qualified, Male})$$

Egalitarian

Equal false positive rates, FPRs

P(Shoot | Unarmed, White) = P(Shoot | Unarmed, Black)

Equal false negative rates, FNRs

P(Undertreat | In Pain, White) = P(Undertreat | In Pain, Black)

Equal positive predictive values, PPVs

P(Hire | Qualified, Female) = P(Hire | Qualified, Male)

Why these three ways of thinking about egalitarian values?

$$FPR_{\text{Group A}} = FPR_{\text{Group B}}$$

$$FNR_{\text{Group A}} = FNR_{\text{Group B}}$$

$$PPV_{\text{Group A}} = PPV_{\text{Group B}}$$

Why these three ways of thinking about egalitarian values?

$$FPR_{\text{Group A}} = FPR_{\text{Group B}}$$

$$FNR_{\text{Group A}} = FNR_{\text{Group B}}$$

$$PPV_{\text{Group A}} = PPV_{\text{Group B}}$$

When base rates between Group A and Group B differ, all three definitions cannot be simultaneously met. At least one must be given up.

Kleinberg, Mullainathan, & Raghavan (2016)

The sinister tradeoff

$$P(\text{cancer} \mid \text{pos}) = \frac{P(\text{pos} \mid \text{cancer}) \times P(\text{cancer})}{P(\text{pos} \mid \text{cancer}) \times P(\text{cancer}) + P(\text{pos} \mid \text{no cancer}) \times P(\text{no cancer})}$$

The sinister tradeoff

$$P(\text{cancer} \mid \text{pos}) = \frac{P(\text{pos} \mid \text{cancer}) \times P(\text{cancer})}{P(\text{pos} \mid \text{cancer}) \times P(\text{cancer}) + P(\text{pos} \mid \text{no cancer}) \times P(\text{no cancer})}$$

$$P(\text{cancer} \mid \text{pos}) = \frac{[1 - P(\text{neg} \mid \text{cancer})] \times P(\text{cancer})}{[1 - P(\text{neg} \mid \text{cancer})] \times P(\text{cancer}) + P(\text{pos} \mid \text{no cancer}) \times [1 - P(\text{cancer})]}$$

The sinister tradeoff

$$P(\text{cancer} \mid \text{pos}) = \frac{P(\text{pos} \mid \text{cancer}) \times P(\text{cancer})}{P(\text{pos} \mid \text{cancer}) \times P(\text{cancer}) + P(\text{pos} \mid \text{no cancer}) \times P(\text{no cancer})}$$

$$P(\text{cancer} \mid \text{pos}) = \frac{[1 - P(\text{neg} \mid \text{cancer})] \times P(\text{cancer})}{[1 - P(\text{neg} \mid \text{cancer})] \times P(\text{cancer}) + P(\text{pos} \mid \text{no cancer}) \times [1 - P(\text{cancer})]}$$

$$\text{PPV} = \frac{[1 - \text{FNR}] \times \text{BR}}{[1 - \text{FNR}] \times \text{BR} + \text{FPR} \times [1 - \text{BR}]}$$

The sinister tradeoff

$$P(\text{cancer} \mid \text{pos}) = \frac{P(\text{pos} \mid \text{cancer}) \times P(\text{cancer})}{P(\text{pos} \mid \text{cancer}) \times P(\text{cancer}) + P(\text{pos} \mid \text{no cancer}) \times P(\text{no cancer})}$$

$$P(\text{cancer} \mid \text{pos}) = \frac{[1 - P(\text{neg} \mid \text{cancer})] \times P(\text{cancer})}{[1 - P(\text{neg} \mid \text{cancer})] \times P(\text{cancer}) + P(\text{pos} \mid \text{no cancer}) \times [1 - P(\text{cancer})]}$$

$$\text{PPV} = \frac{[1 - \text{FNR}] \times \text{BR}}{[1 - \text{FNR}] \times \text{BR} + \text{FPR} \times [1 - \text{BR}]}$$

$$\text{FPR} = \frac{\text{BR}}{[1 - \text{BR}]} \times \frac{[1 - \text{PPV}]}{\text{PPV}} \times [1 - \text{FNR}]$$

The sinister tradeoff

$$FPR_w = \frac{BR_w}{1 - BR_w} \times \frac{1 - PPV_w}{PPV_w} \times [1 - FNR_w]$$

$$FPR_m = \frac{BR_m}{1 - BR_m} \times \frac{1 - PPV_m}{PPV_m} \times [1 - FNR_m]$$

The sinister tradeoff

$$FPR_w = \frac{BR_w}{1-BR_w} \times \frac{1-PPV_w}{PPV_w} \times [1-FNR_w]$$

$$FPR_m = \frac{BR_m}{1-BR_m} \times \frac{1-PPV_m}{PPV_m} \times [1-FNR_m]$$

Meet two definitions of fairness

$PPV_w = PPV_m$

$FNR_w = FNR_m$

$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w} \times \frac{1-PPV_w}{PPV_w} \times [1-FNR_w]}{\frac{BR_m}{1-BR_m} \times \frac{1-PPV_m}{PPV_m} \times [1-FNR_m]}$$

The sinister tradeoff

$$FPR_w = \frac{BR_w}{1-BR_w} \times \frac{1-PPV_w}{PPV_w} \times [1-FNR_w]$$

$$FPR_m = \frac{BR_m}{1-BR_m} \times \frac{1-PPV_m}{PPV_m} \times [1-FNR_m]$$

Meet two definitions of fairness

$PPV_w = PPV_m$

$FNR_w = FNR_m$

$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w} \times \cancel{\frac{1-PPV_w}{PPV_w}} \times \cancel{[1-FNR_w]}}{\frac{BR_m}{1-BR_m} \times \cancel{\frac{1-PPV_m}{PPV_m}} \times \cancel{[1-FNR_m]}}$$

The sinister tradeoff

Meet two definitions of fairness

$$PPV_w = PPV_m$$

$$FNR_w = FNR_m$$

$$FPR_w = \frac{BR_w}{1-BR_w} \times \frac{1-PPV_w}{PPV_w} \times [1-FNR_w]$$

$$FPR_m = \frac{BR_m}{1-BR_m} \times \frac{1-PPV_m}{PPV_m} \times [1-FNR_m]$$

$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w} \times \cancel{\frac{1-PPV_w}{PPV_w}} \times \cancel{[1-FNR_w]}}{\frac{BR_m}{1-BR_m} \times \cancel{\frac{1-PPV_m}{PPV_m}} \times \cancel{[1-FNR_m]}}$$

$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w}}{\frac{BR_m}{1-BR_m}}$$

The sinister tradeoff

Meet two definitions of fairness

$$PPV_w = PPV_m$$

$$FNR_w = FNR_m$$

$$FPR_w = \frac{BR_w}{1-BR_w} \times \frac{1-PPV_w}{PPV_w} \times [1-FNR_w]$$

$$FPR_m = \frac{BR_m}{1-BR_m} \times \frac{1-PPV_m}{PPV_m} \times [1-FNR_m]$$

$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w} \times \cancel{\frac{1-PPV_w}{PPV_w}} \times \cancel{[1-FNR_w]}}{\frac{BR_m}{1-BR_m} \times \cancel{\frac{1-PPV_m}{PPV_m}} \times \cancel{[1-FNR_m]}}$$

$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w}}{\frac{BR_m}{1-BR_m}}$$

$$FPR_w = \frac{BR_w}{BR_m} \times \frac{[1-BR_m]}{[1-BR_w]} \times FPR_m$$

The sinister tradeoff

$$FPR_w = \frac{BR_w}{1-BR_w} \times \frac{1-PPV_w}{PPV_w} \times [1-FNR_w]$$

$$FPR_m = \frac{BR_m}{1-BR_m} \times \frac{1-PPV_m}{PPV_m} \times [1-FNR_m]$$

Meet two definitions of fairness

$PPV_w = PPV_m$

$FNR_w = FNR_m$

$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w} \times \cancel{\frac{1-PPV_w}{PPV_w}} \times \cancel{[1-FNR_w]}}{\frac{BR_m}{1-BR_m} \times \cancel{\frac{1-PPV_m}{PPV_m}} \times \cancel{[1-FNR_m]}}$$

$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w}}{\frac{BR_m}{1-BR_m}}$$

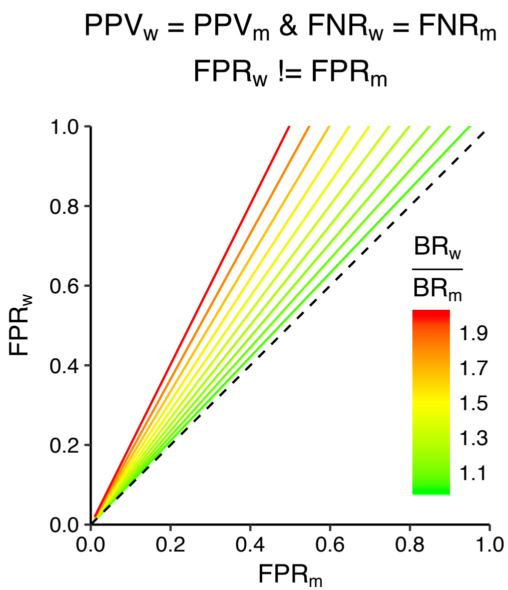
Cannot meet third definition of fairness

$FPR_w \;!= FPR_m$

Breast cancer is more common among women than men

$$FPR_w = \frac{BR_w}{BR_m} \times \frac{[1-BR_m]}{[1-BR_w]} \times FPR_m$$

The sinister tradeoff



$$FPR_w = \frac{BR_w}{1-BR_w} \times \frac{1-PPV_w}{PPV_w} \times [1-FNR_w]$$

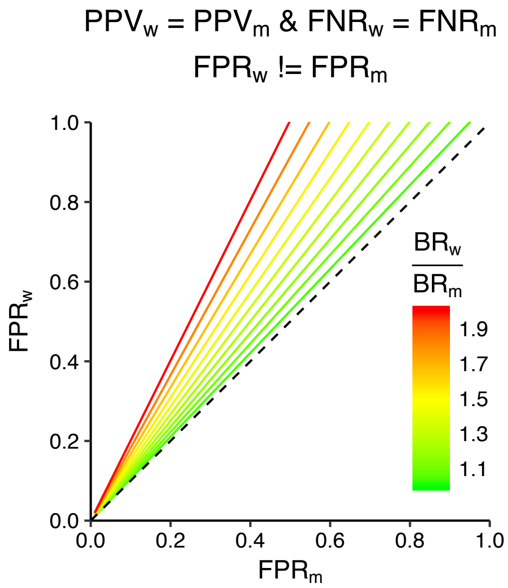
$$FPR_m = \frac{BR_m}{1-BR_m} \times \frac{1-PPV_m}{PPV_m} \times [1-FNR_m]$$

$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w} \times \cancel{\frac{1-PPV_w}{PPV_w}} \times \cancel{[1-FNR_w]}}{\frac{BR_m}{1-BR_m} \times \cancel{\frac{1-PPV_m}{PPV_m}} \times \cancel{[1-FNR_m]}}$$

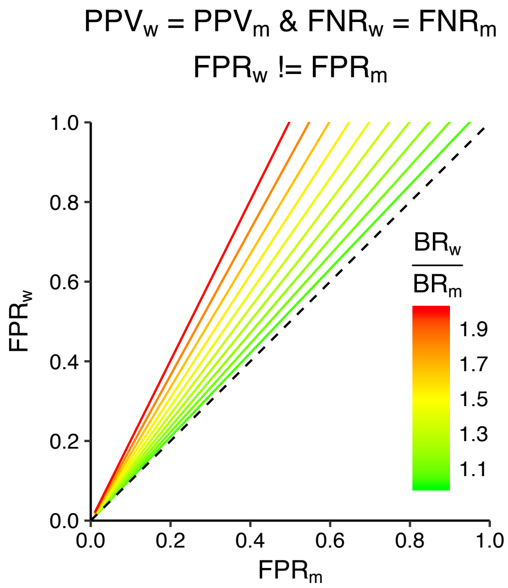
$$\frac{FPR_w}{FPR_m} = \frac{\frac{BR_w}{1-BR_w}}{\frac{BR_m}{1-BR_m}}$$

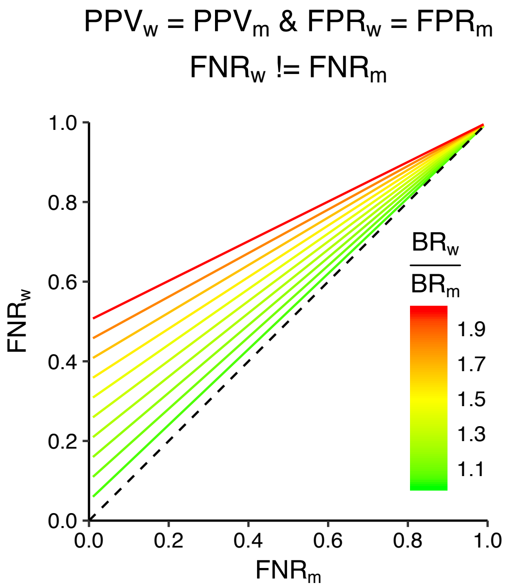
$$FPR_w = \frac{BR_w}{BR_m} \times \frac{[1-BR_m]}{[1-BR_w]} \times FPR_m$$

The sinister tradeoff

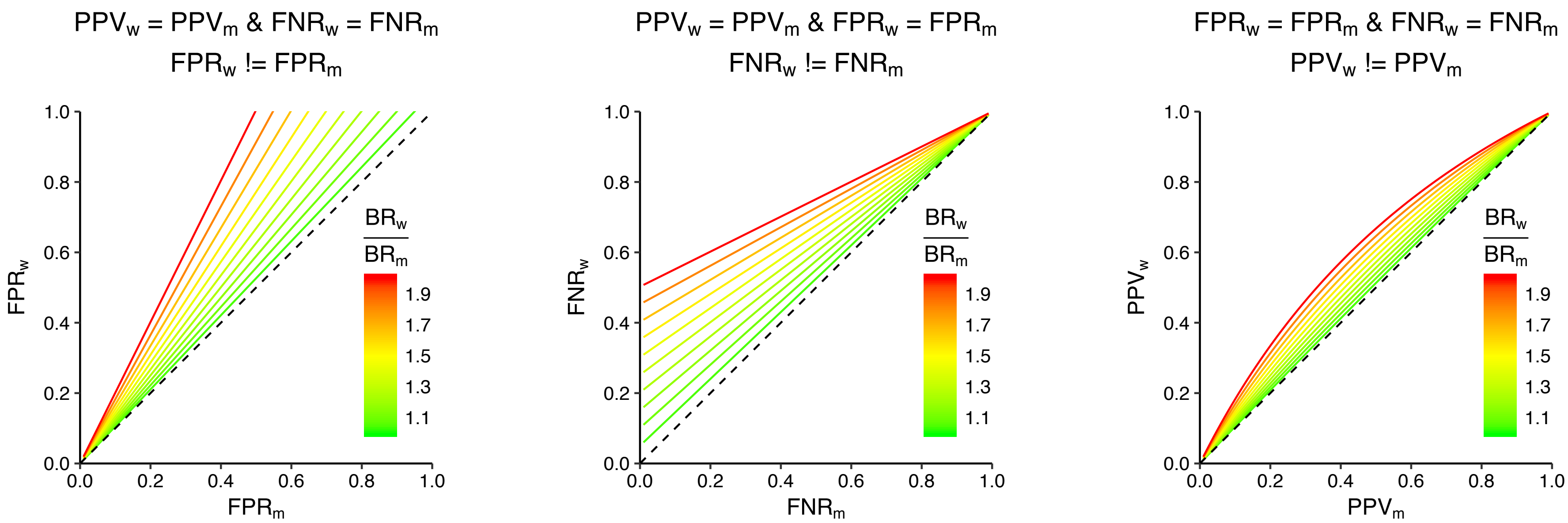


The sinister tradeoff

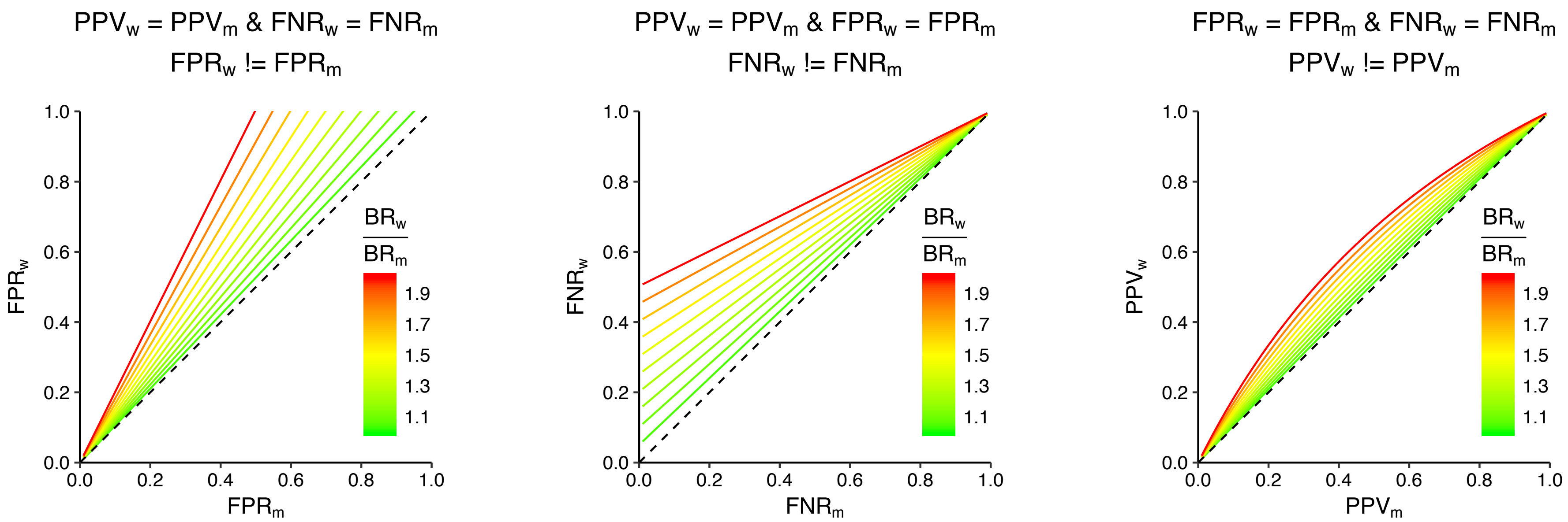




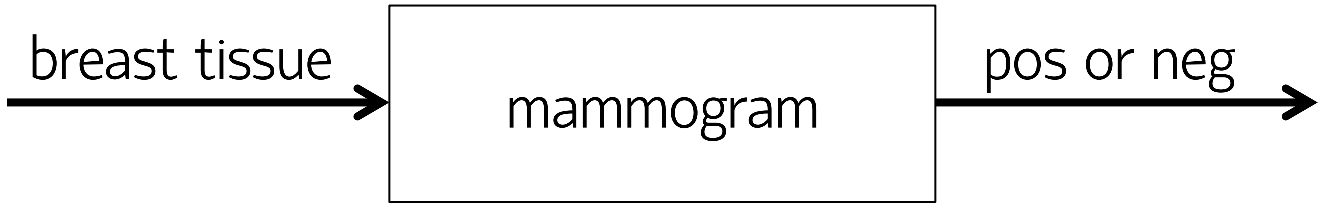
The sinister tradeoff



The sinister tradeoff



Of the three "worlds" depicted above, which is preferred?
Which notion of fairness is valued more? Which is valued less?



Bail eligibility for defendants

Bank loans

Whether a post is fake news

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

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

Machine Bias

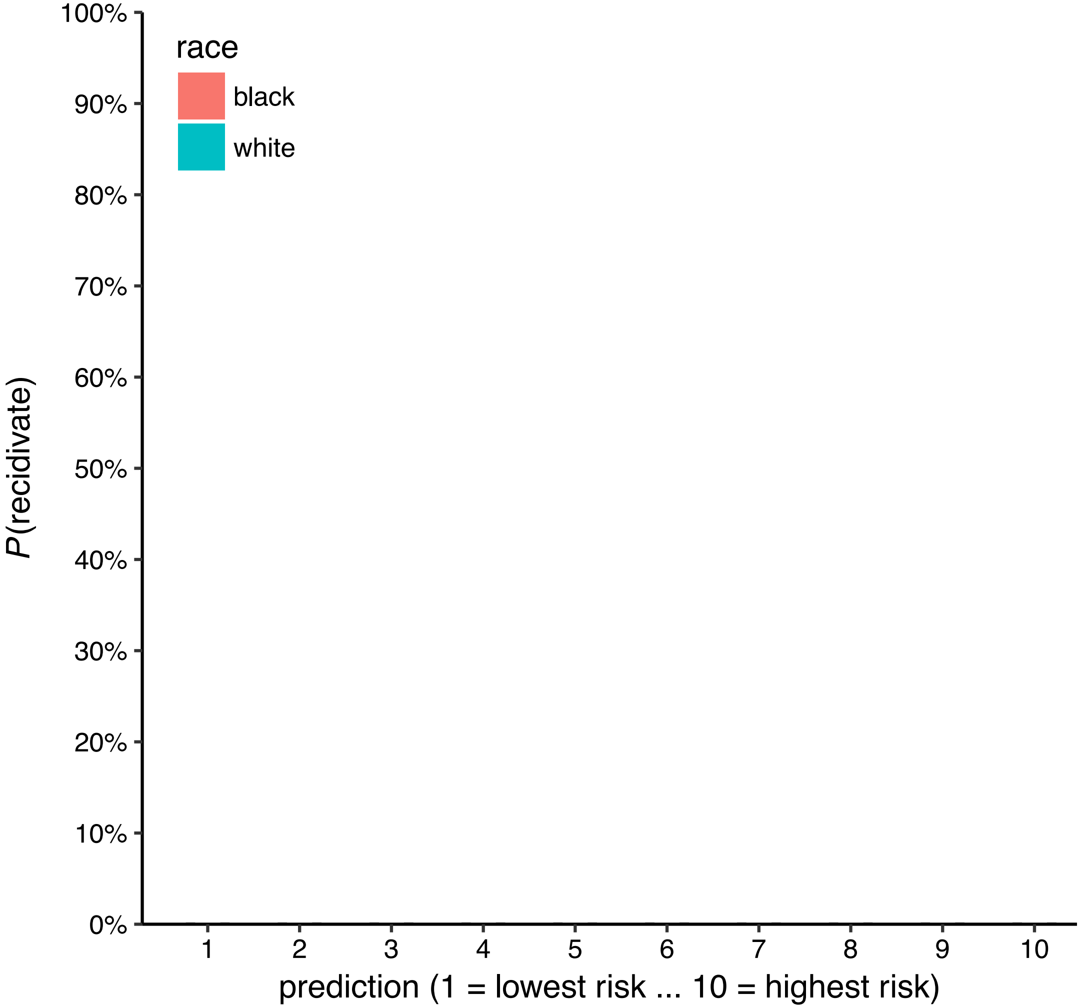
There's software used across the country to predict future criminals. And it's biased against blacks.

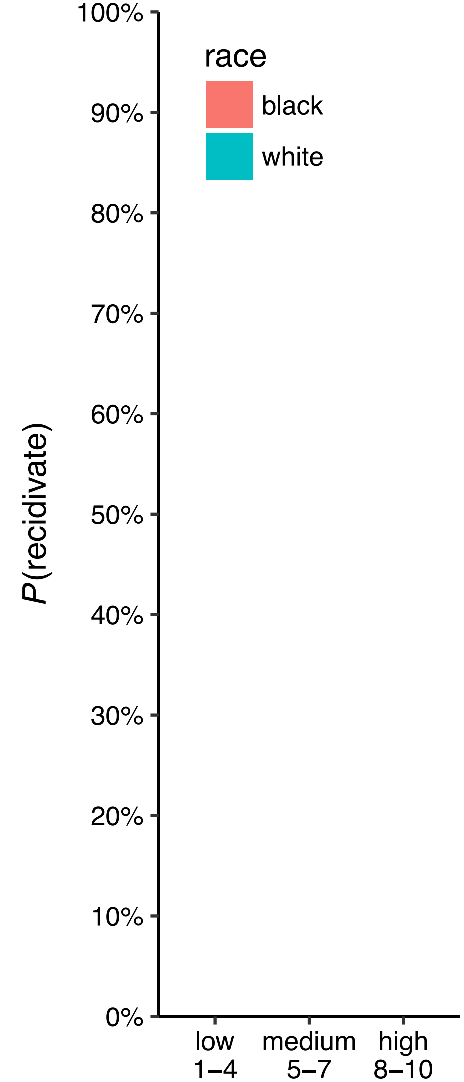
False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks."

Anthony W. Flores
California State University, Bakersfield
Kristin Bechtel
Crime and Justice Institute at CRJ
Christopher T. Lowenkamp
Administrative Office of the United States Courts
Probation and Pretrial Services Office

$$\mathrm{PPV}_{\mathrm{black}} = \mathrm{PPV}_{\mathrm{white}}$$

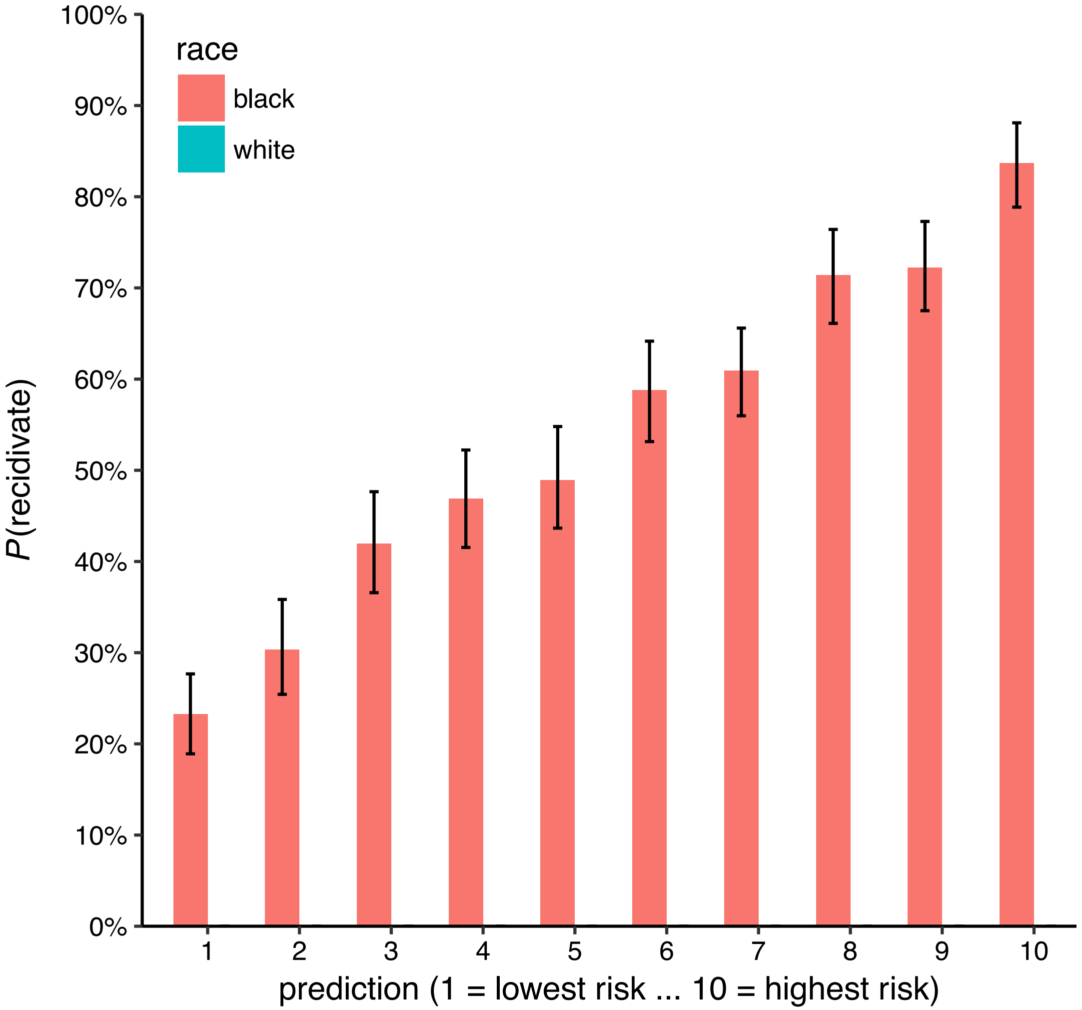
$$P(\text{recidivate} \mid \text{risk score, black}) = P(\text{recidivate} \mid \text{risk score, white})$$

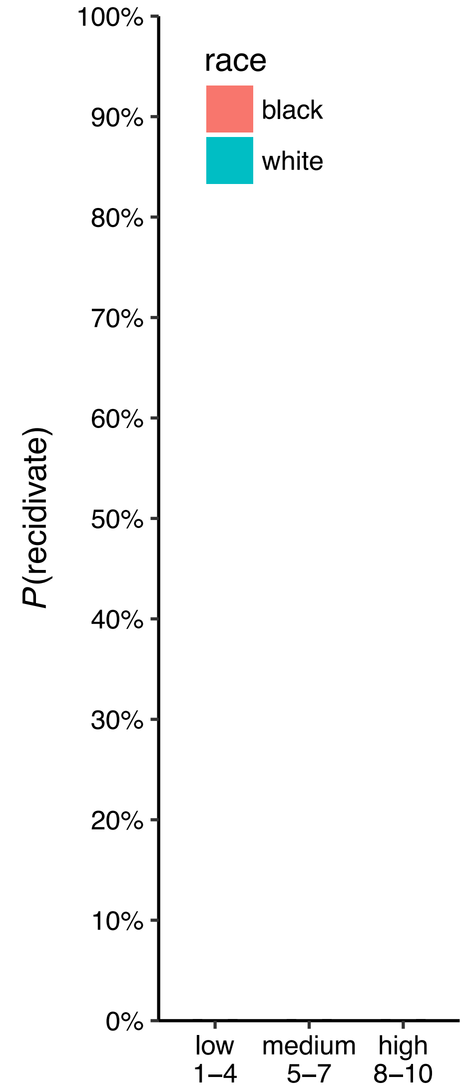




$$PPV_{black} = PPV_{white}$$

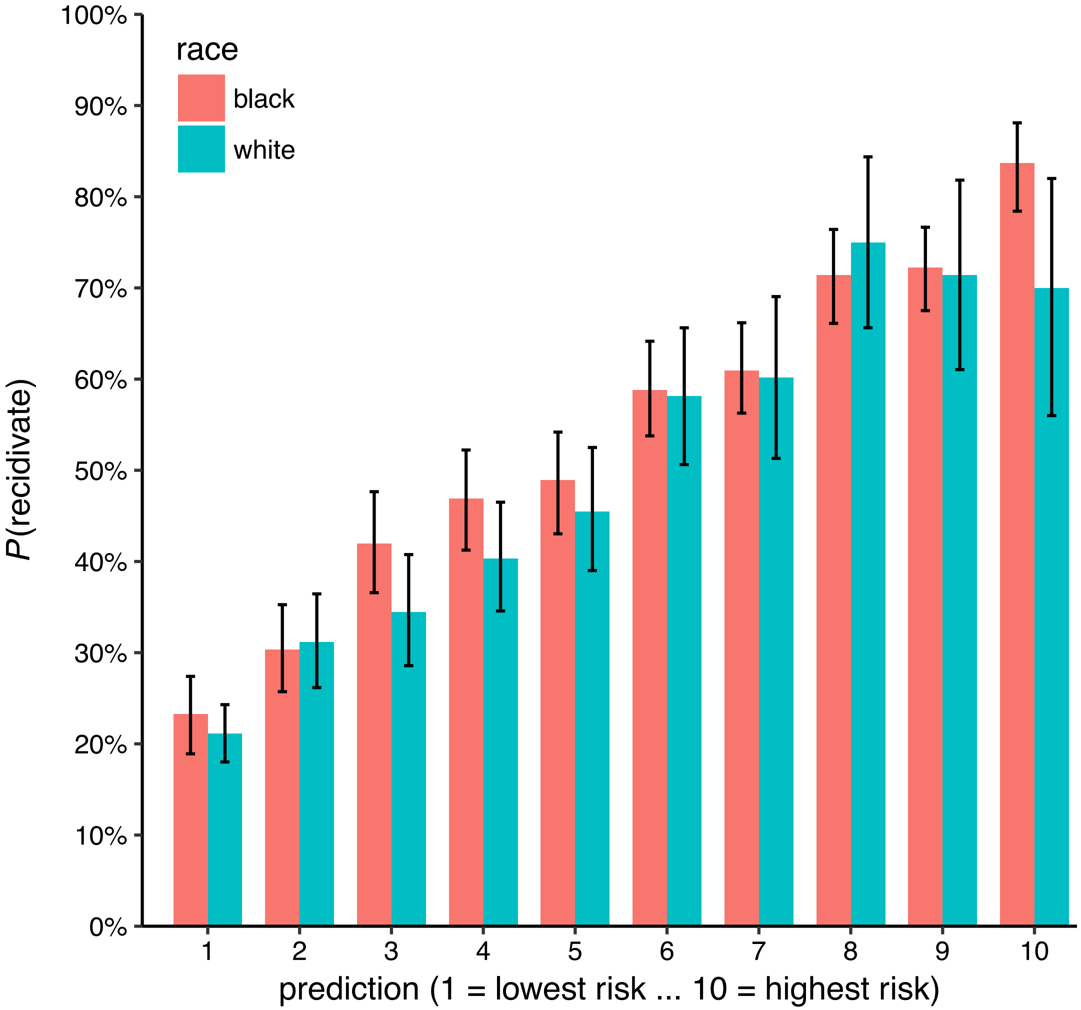
$$P(\text{recidivate} \mid \text{risk score, black}) = P(\text{recidivate} \mid \text{risk score, white})$$

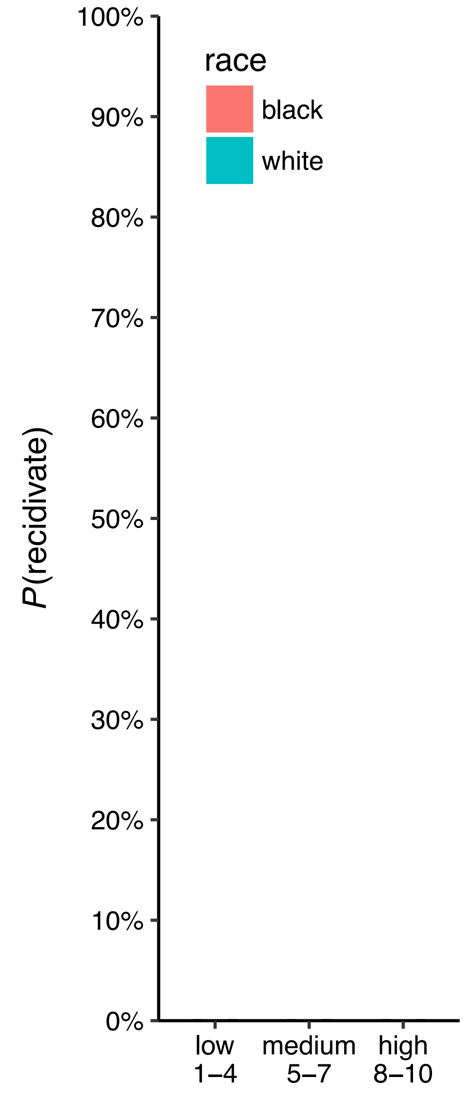




$$PPV_{black} = PPV_{white}$$

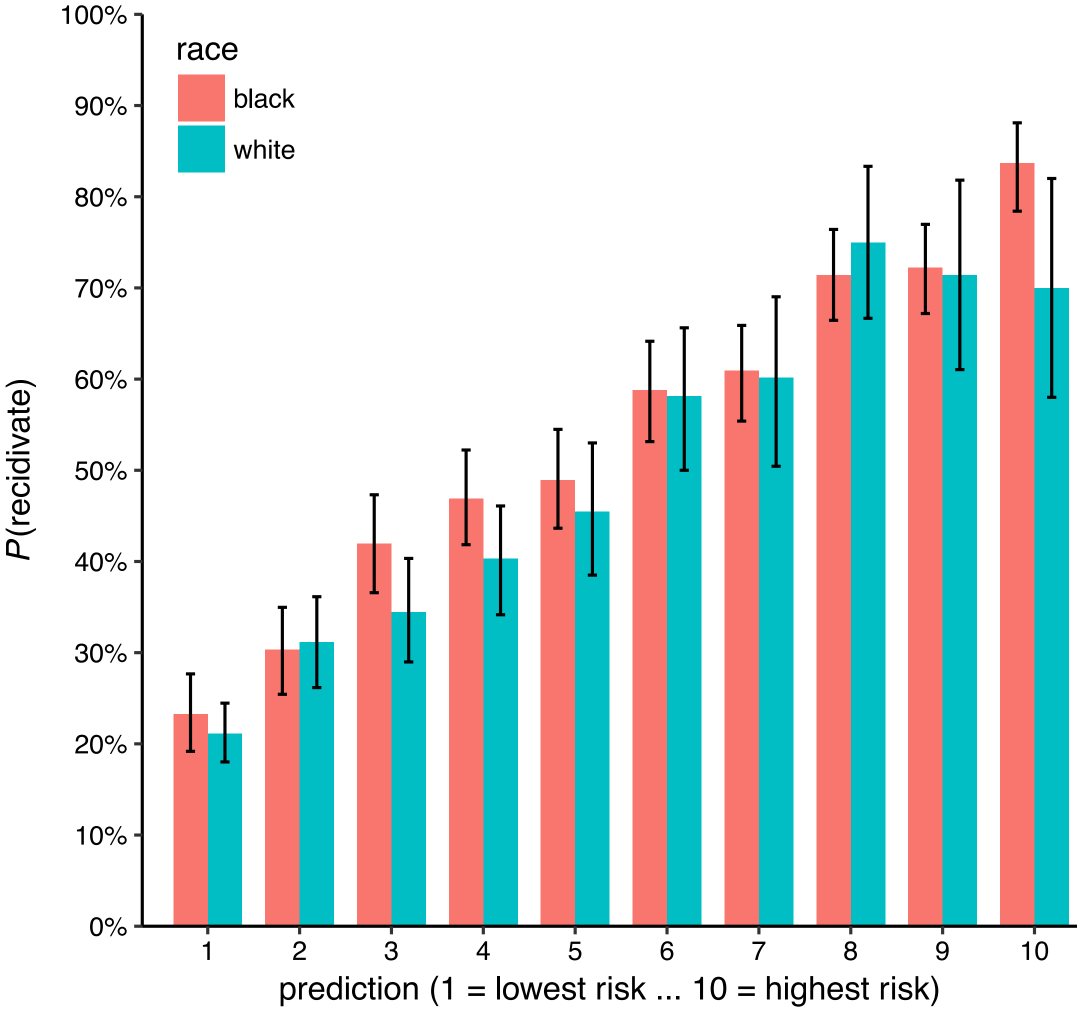
$$P(\text{recidivate} \mid \text{risk score, black}) = P(\text{recidivate} \mid \text{risk score, white})$$

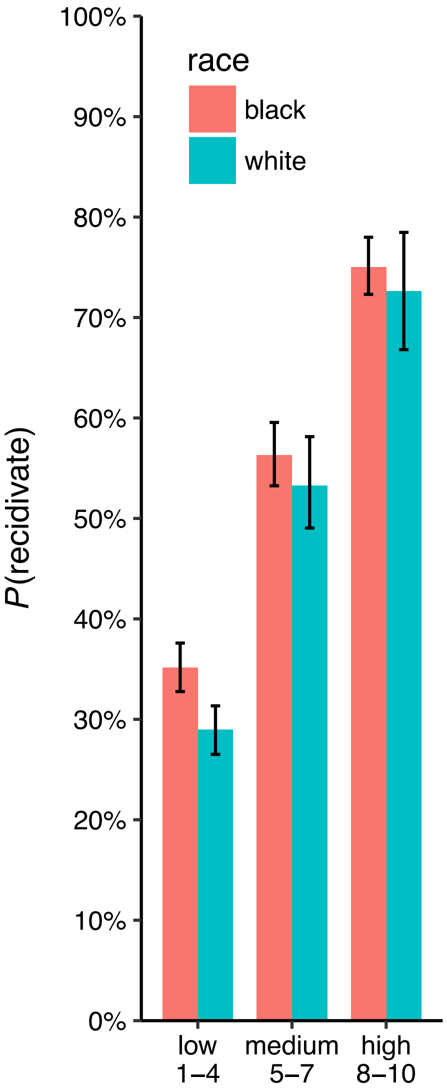




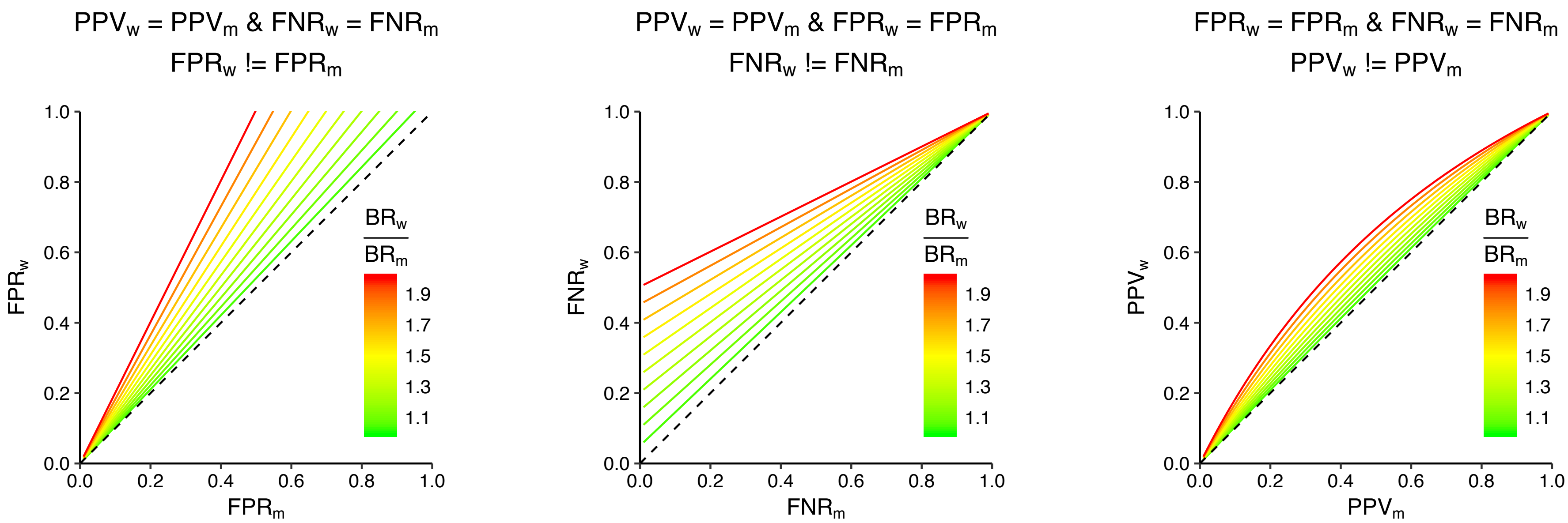
$$PPV_{black} = PPV_{white}$$

$$P(\text{recidivate} \mid \text{risk score, black}) = P(\text{recidivate} \mid \text{risk score, white})$$





The sinister tradeoff



Of the three "worlds" depicted above, which is preferred?
Which notion of fairness is valued more? Which is valued less?

What judgments do people make when Bayesian principles and egalitarian values are at stake?

1. People undermine their commitment to egalitarian values by making Bayesian judgments

2. Formalizing "egalitarian values"

3. Google Images as a proxy for the social environment

What judgments do people make when Bayesian principles and egalitarian values are at stake?

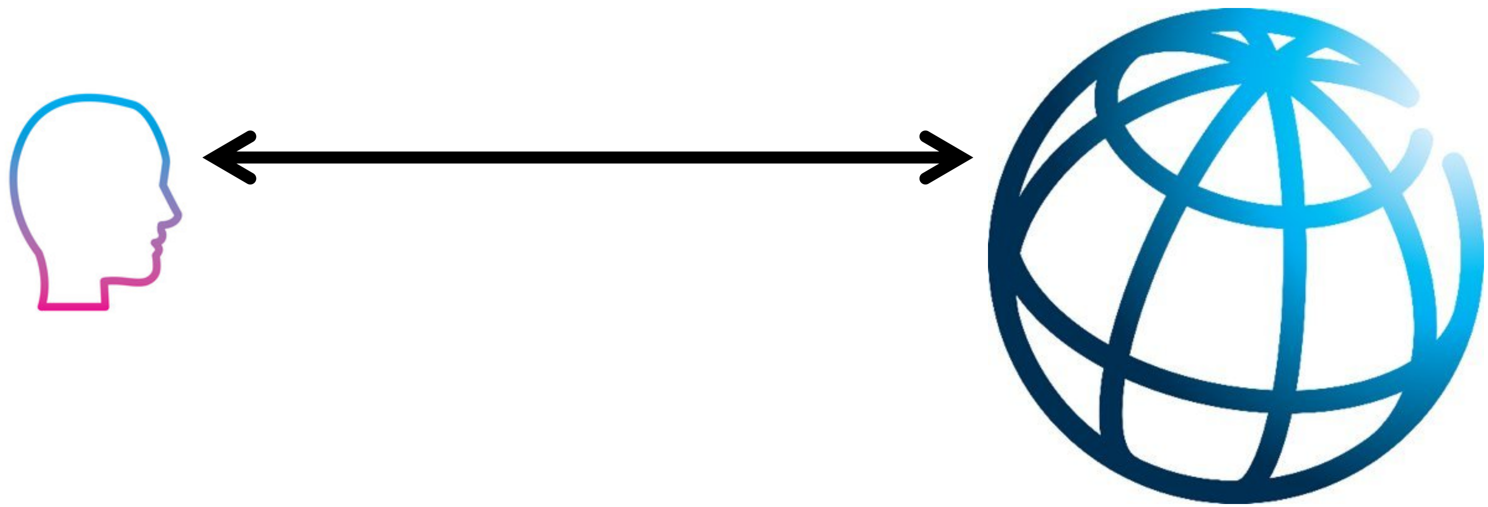
1. People undermine their commitment to egalitarian values by making Bayesian judgments

2. Formalizing "egalitarian values"

3. Google Images as a proxy for the social environment

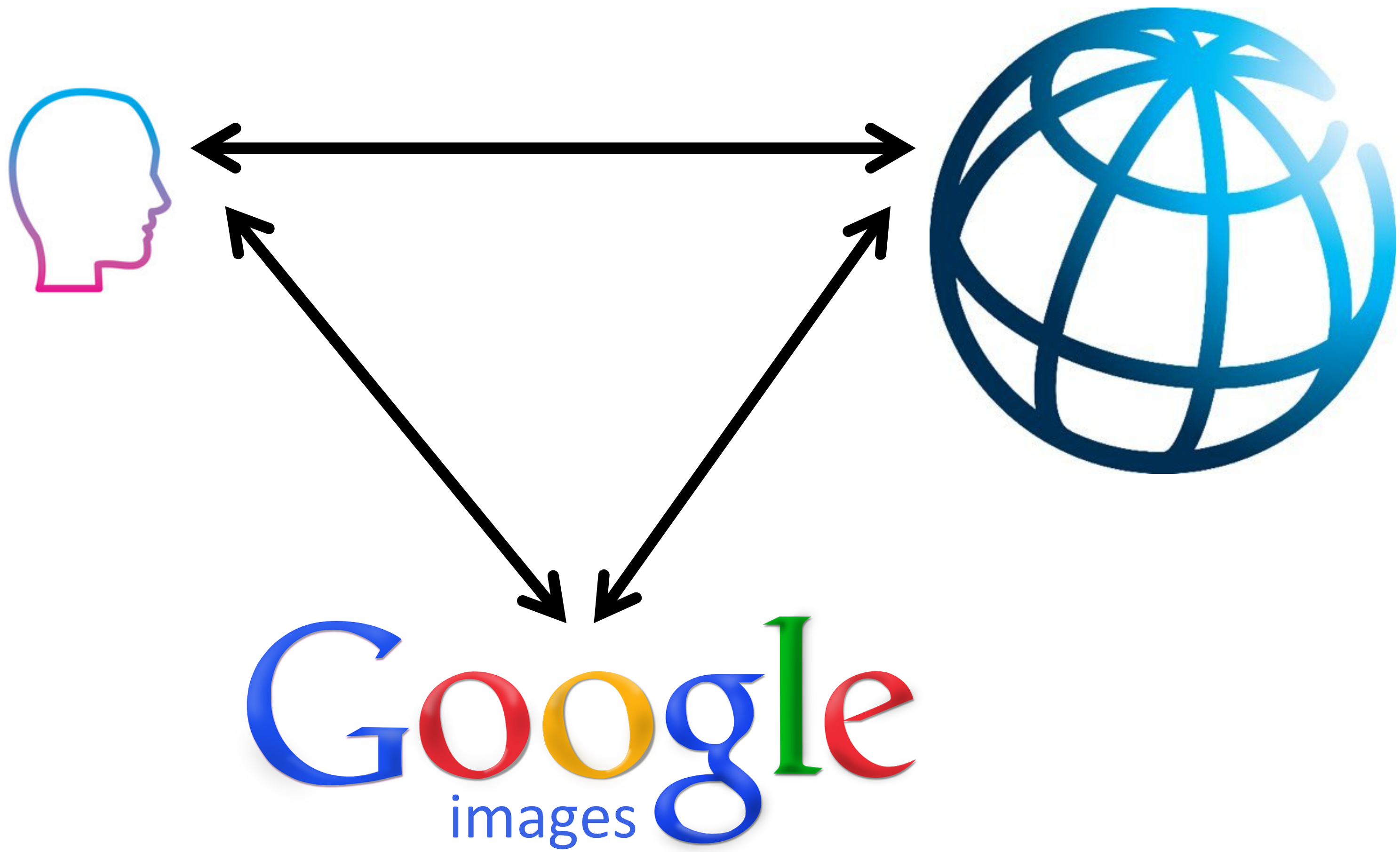
1. Pictures of people matter a great deal

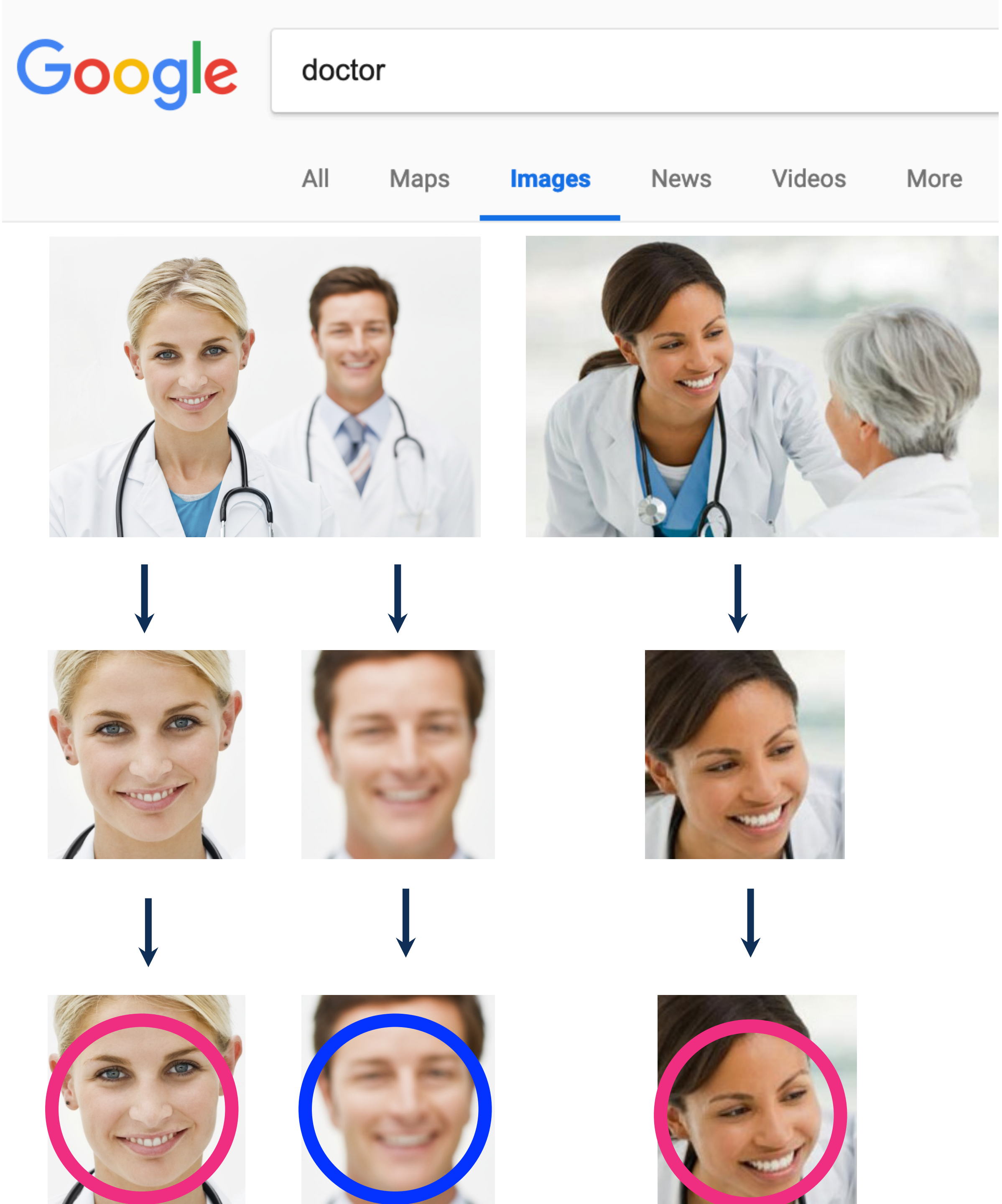
2. There's some relationship between the content in our minds and the content in the world



1. Pictures of people matter a great deal

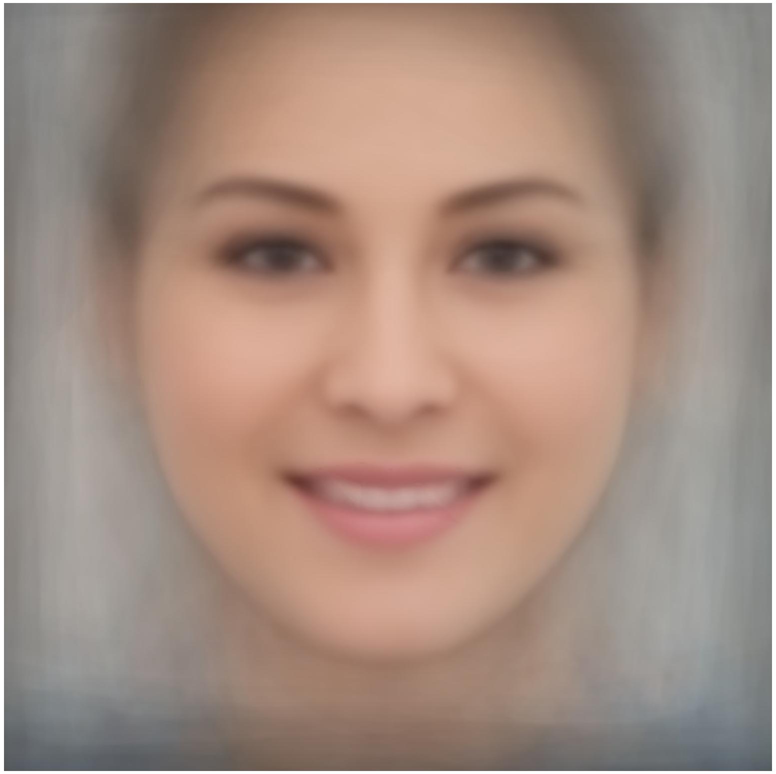
2. There's some relationship between the content in our minds and the content in the world




doctor
All
Maps
Images
News
Videos
More

"doctor"

"nurse"



"computer science student"

"biology student"

"philosophy student"

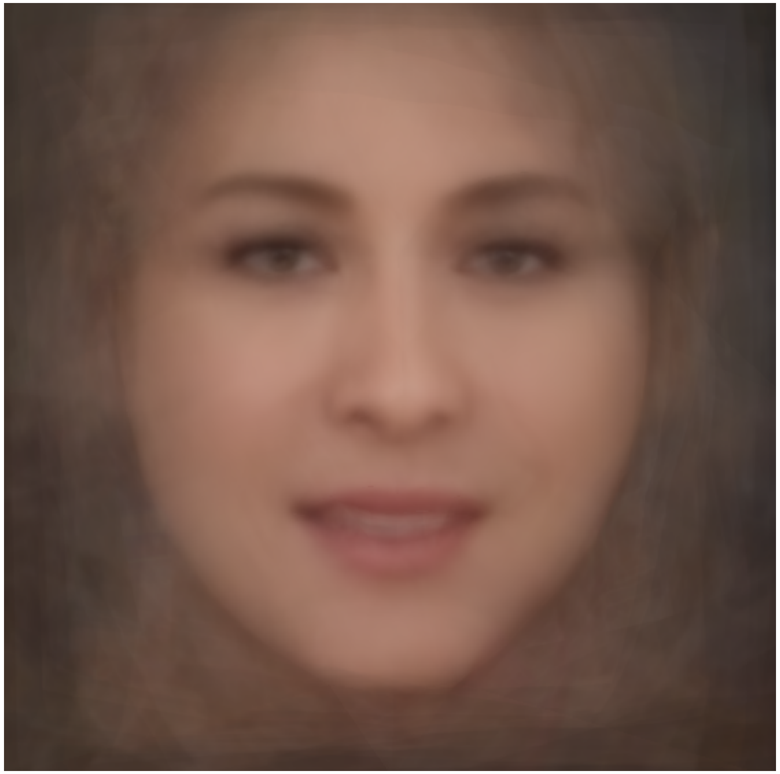
"education student"

"intelligent person"

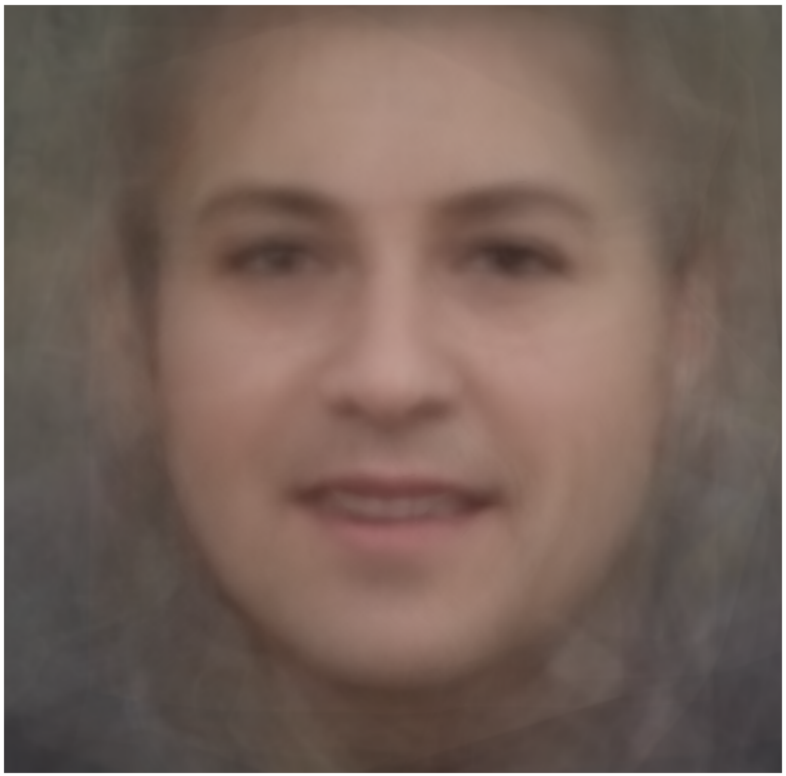
"sensitive person"

"person drinking whiskey"

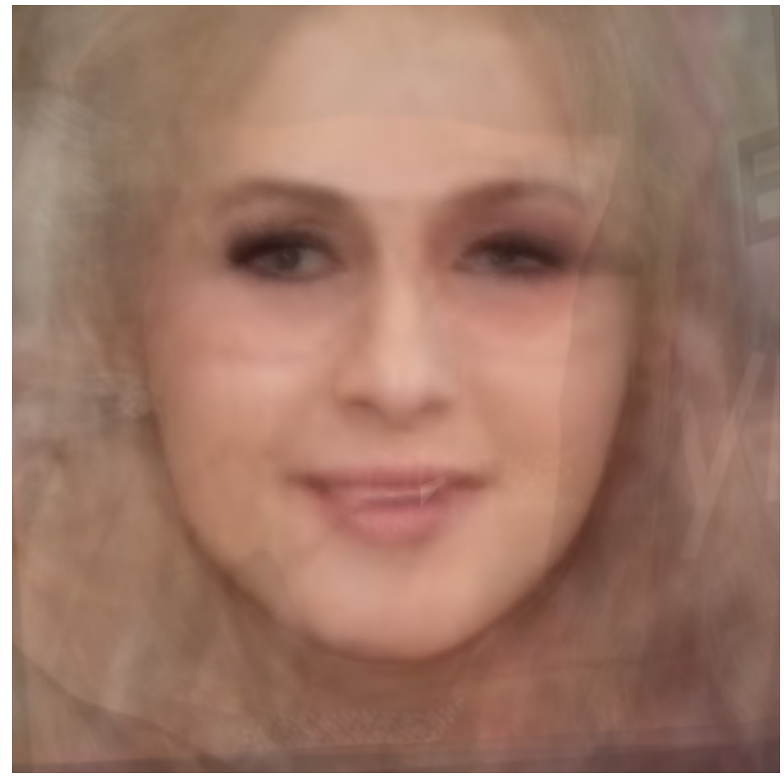
"person drinking cosmopolitan"



"bulldog owner"

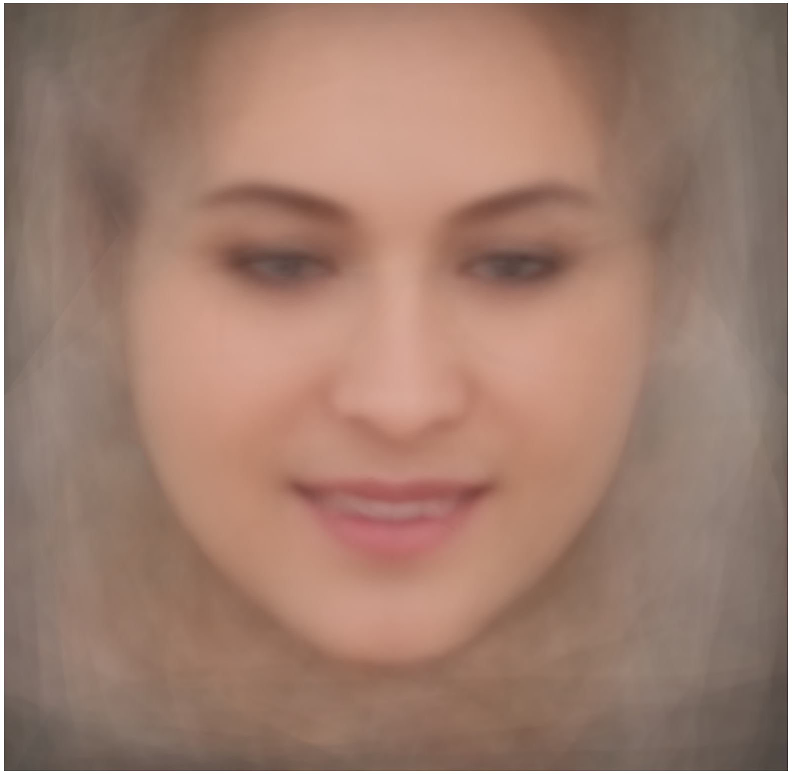


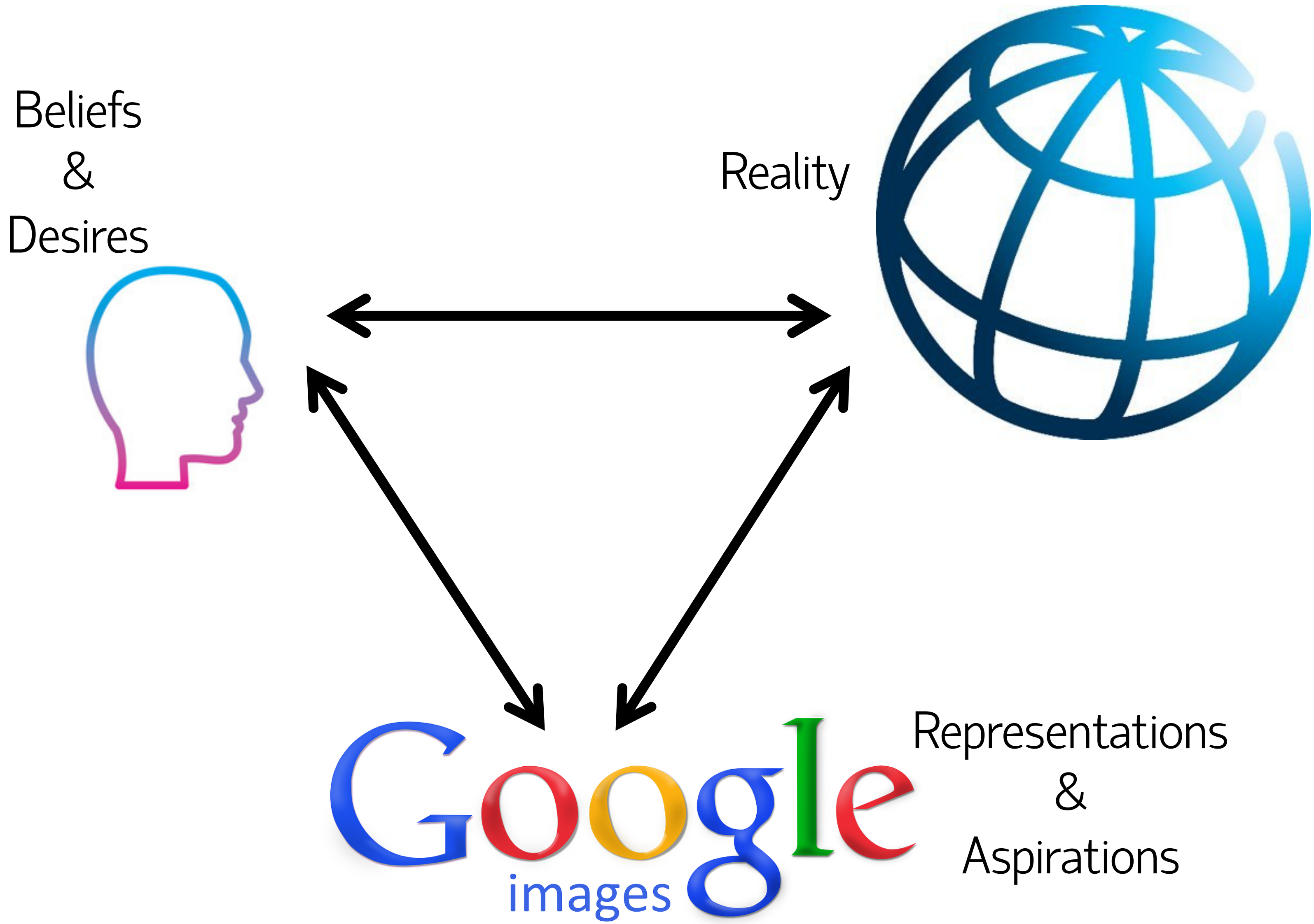
"chihuahua owner"

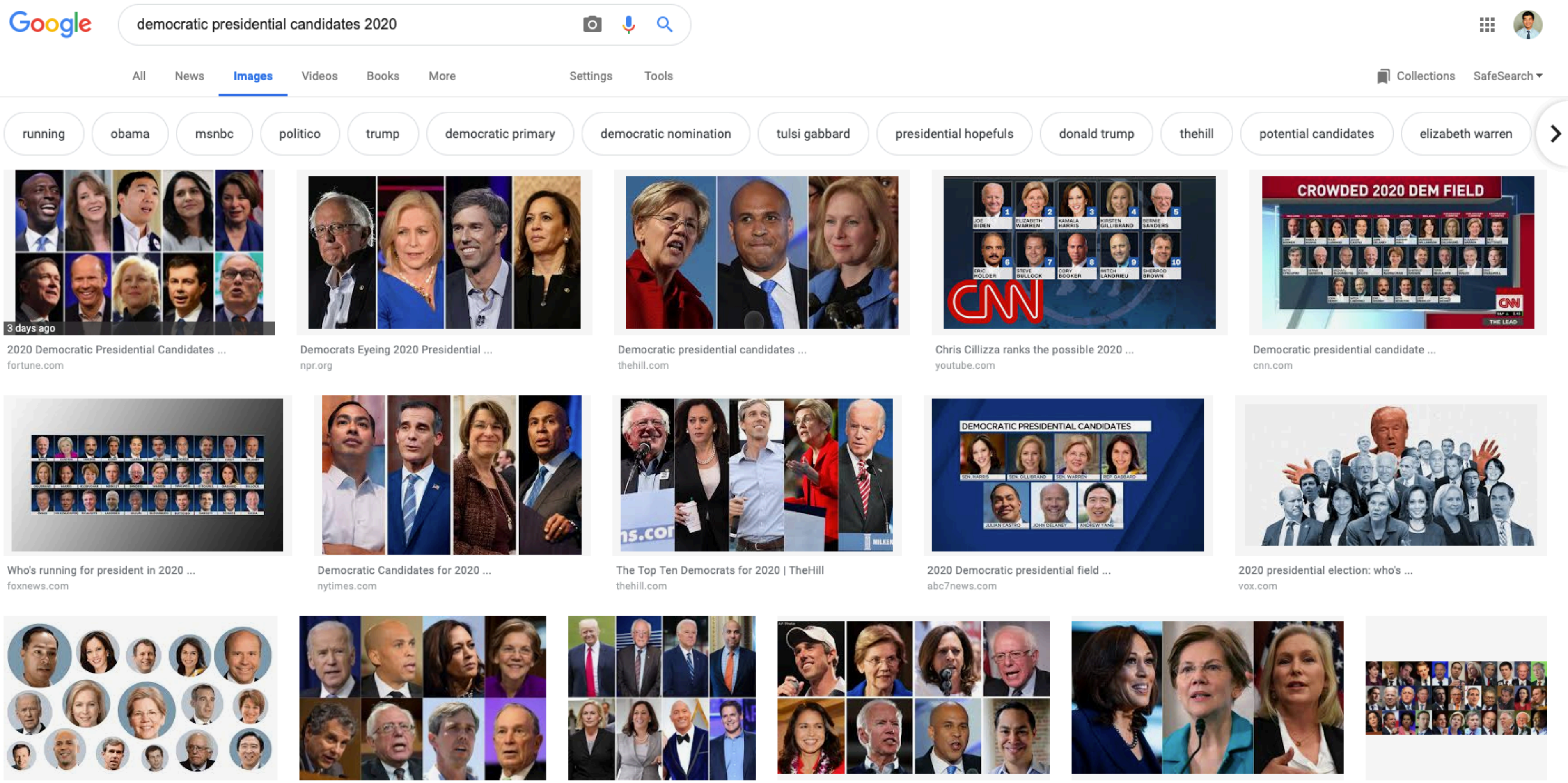


"person reading a newspaper"

"person reading a cookbook"




Beliefs
&
Desires
Reality
Google
images
Representations
&
Aspirations


Google
democratic presidential candidates 2020
All
News
Images
Videos
Books
More
Settings
Tools
Collections
SafeSearch
running
obama
msnbc
politico
trump
democratic primary
democratic nomination
tulsi gabbard
presidential hopefuls
donald trump
thehill
potential candidates
elizabeth warren
3 days ago
2020 Democratic Presidential Candidates ...
fortune.com
Democrats Eyeing 2020 Presidential ...
npr.org
Democratic presidential candidates ...
thehill.com
Chris Cillizza ranks the possible 2020 ...
youtube.com
CROWDED 2020 DEM FIELD
Democratic presidential candidate ...
cnn.com
Who's running for president in 2020 ...
foxnews.com
Democratic Candidates for 2020 ...
nytimes.com
The Top Ten Democrats for 2020 | TheHill
thehill.com
DEMOCRATIC PRESIDENTIAL CANDIDATES
2020 Democratic presidential field ...
abc7news.com
2020 presidential election: who's ...
vox.com

Summary

1. People undermine their commitment to egalitarian values by making Bayesian judgments

2. Formalizing "egalitarian values"

3. Google Images as a proxy for the social environment

Closing thoughts

Life is mostly between-subjects.
Make it more within-subjects.

Need for theory on specifying egalitarian principles.

Relatively low-cost tweaks might go a long way towards promoting inclusion.

Thank you

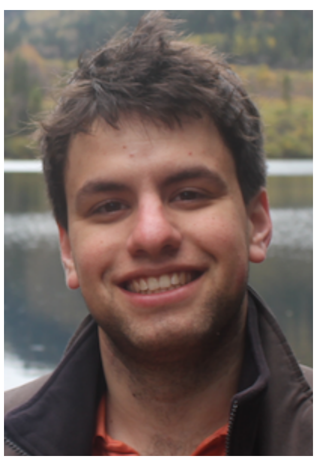
Photo Credit: Robert Taylor www.taylor-photo.co.uk

Mahzarin R. Banaji

Harvard Social Cognition Lab

Jason Mitchell & Jim Sidanius

Research Assistants
Kirsten Morehouse
Juan Lopez Martin



Max Kleiman-Weiner

GRFP
NSF GRADUATE RESEARCH FELLOWSHIP PROGRAM

VE RI TAS
HARVARD
Mind Brain Behavior

HARVARD Kennedy School
MALCOLM WIENER CENTER
for Social Policy