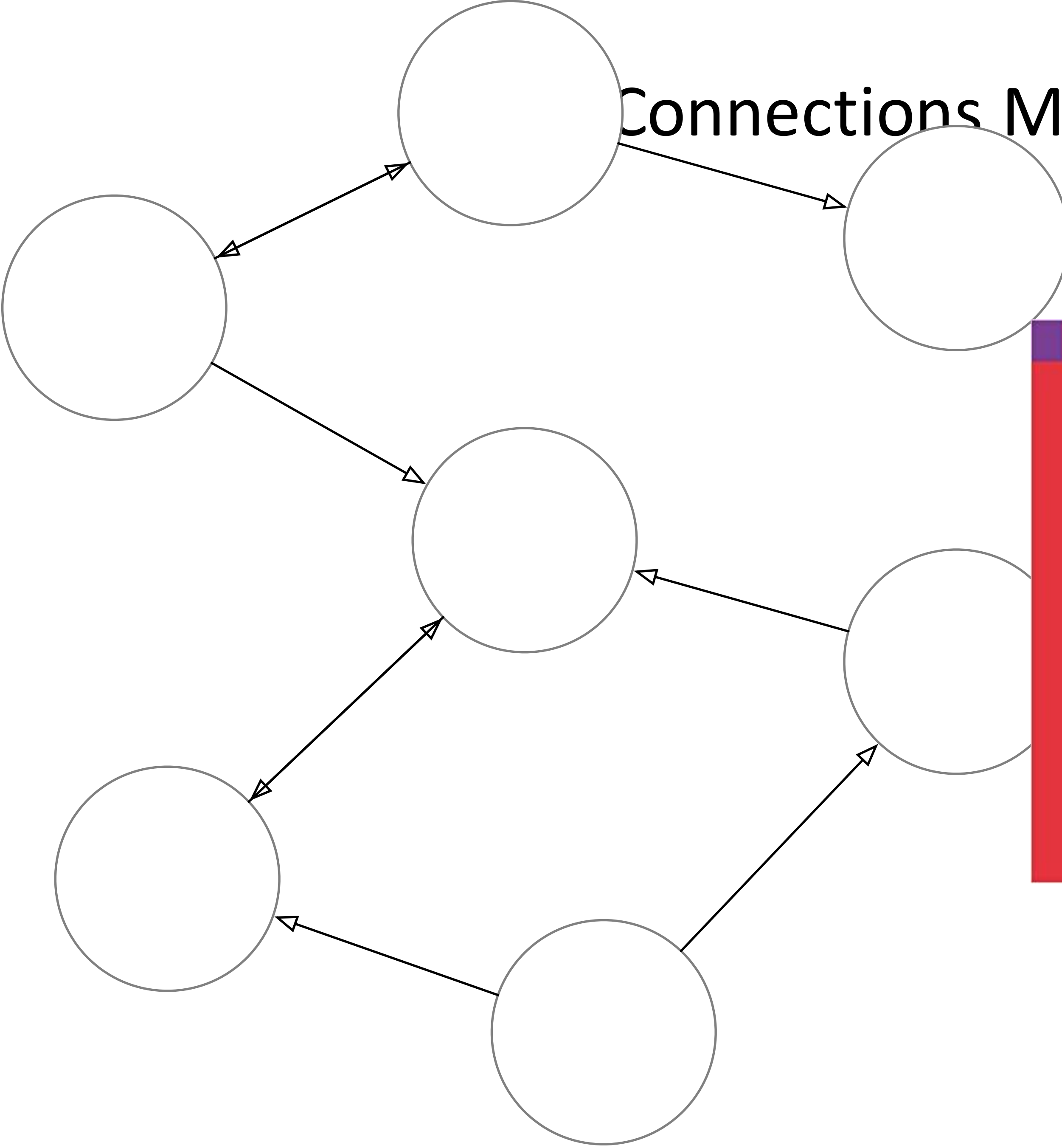
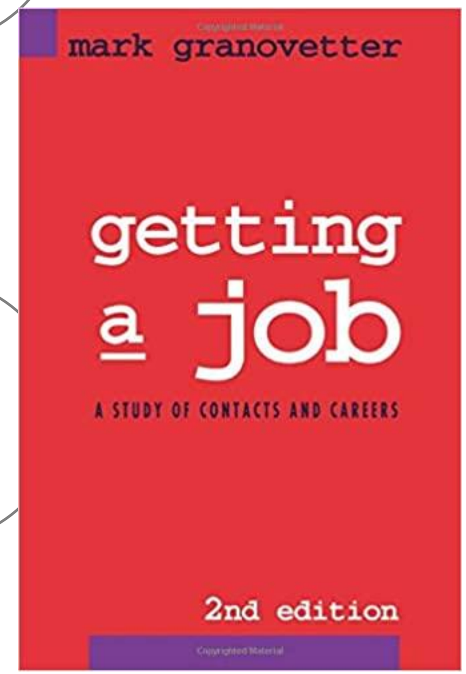
Human Connections and Algorithmic Biases

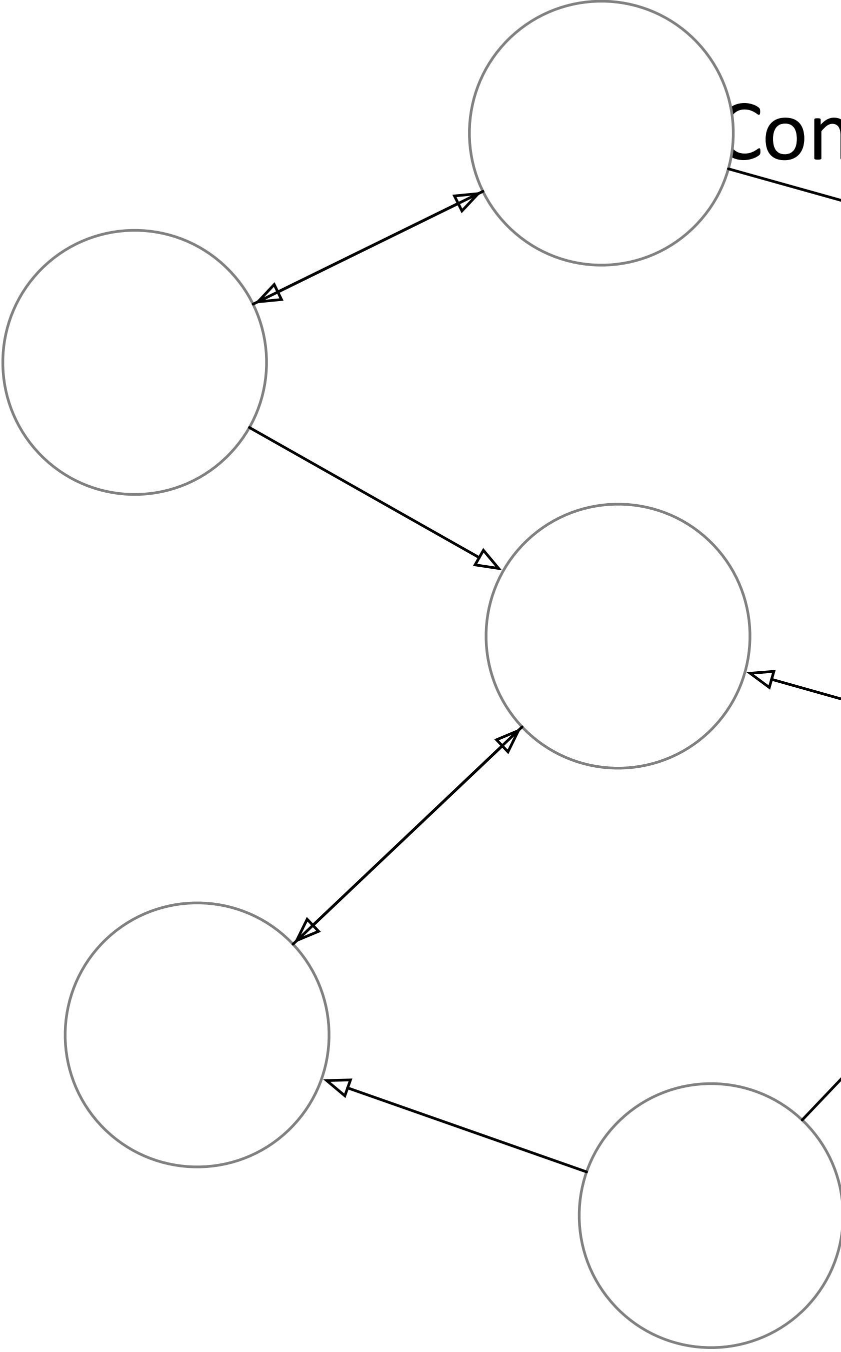
Amanda Agan
Diag Davenport
Jens Ludwig
Sendhil Mullainathan

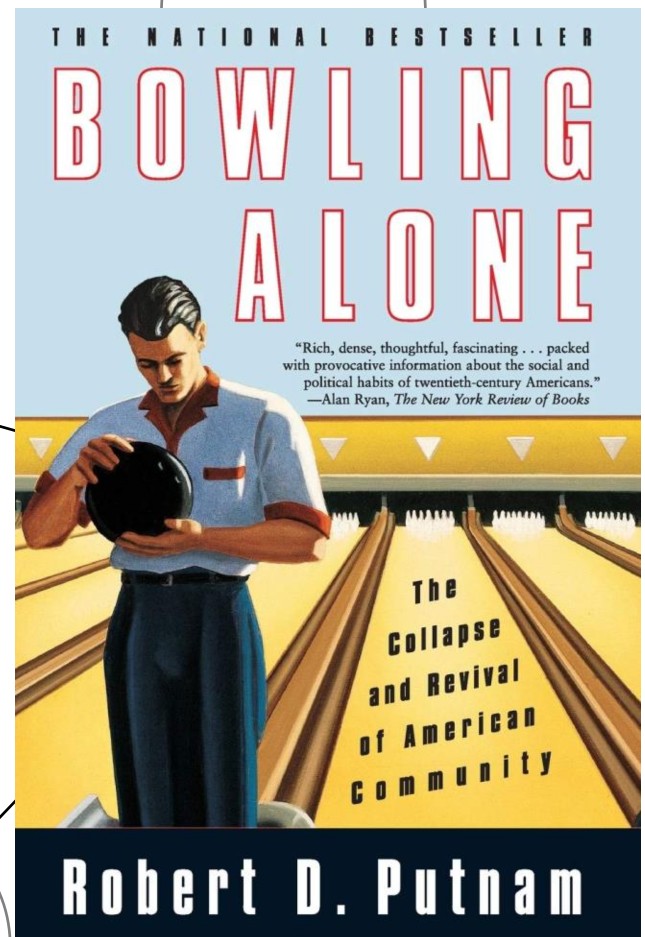
Connections Matter

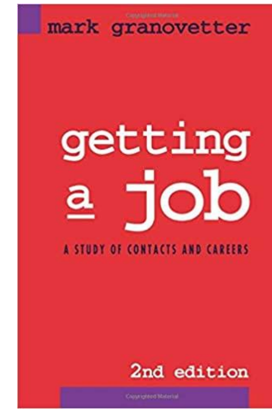




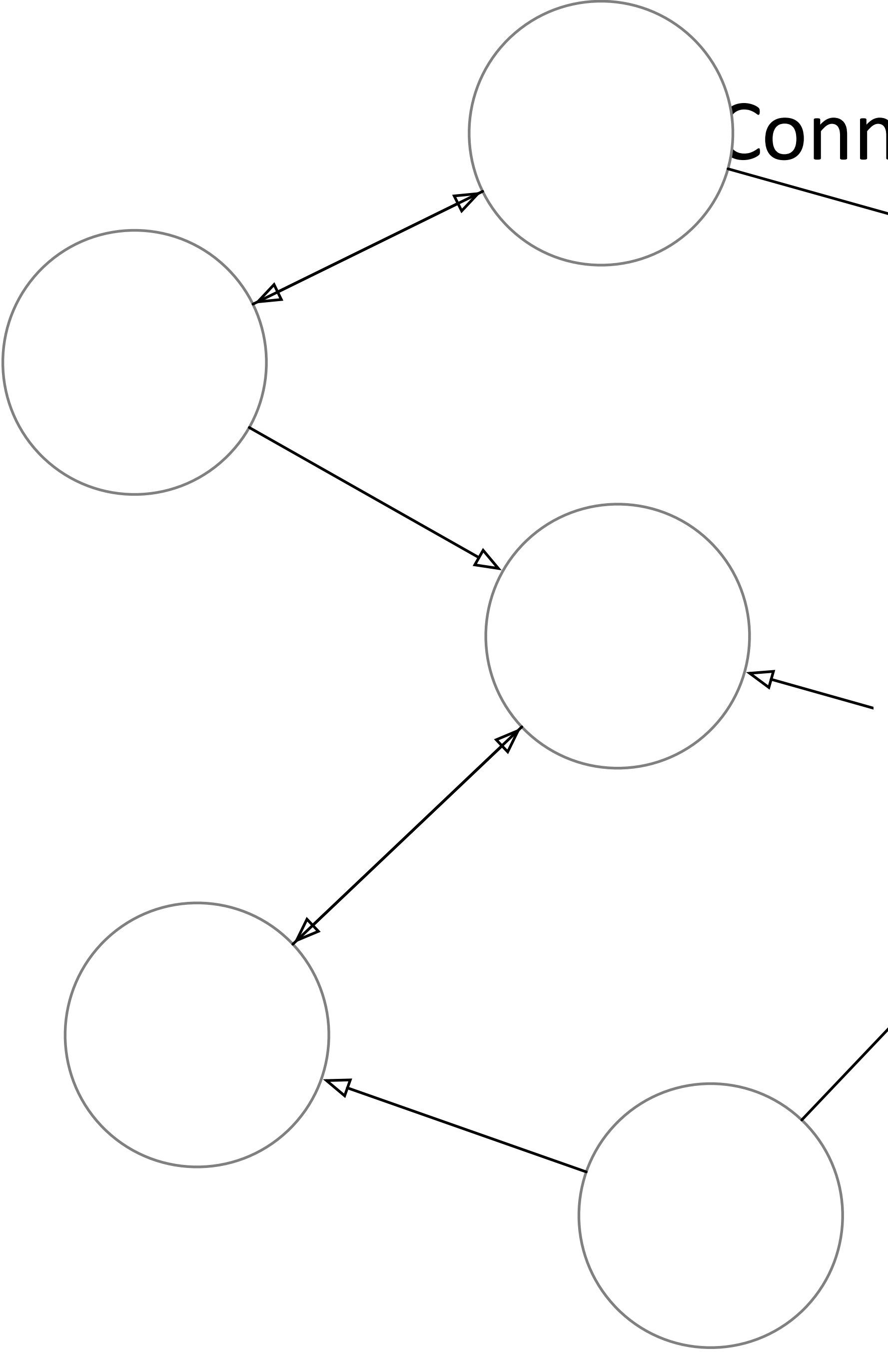
Connections Matter



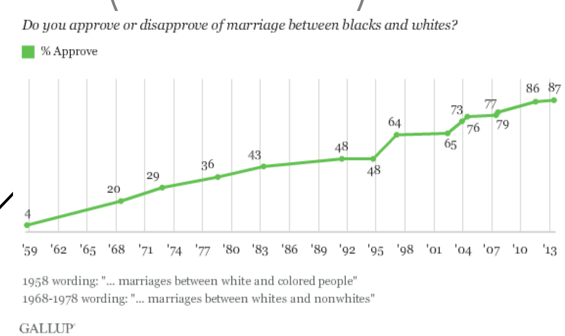


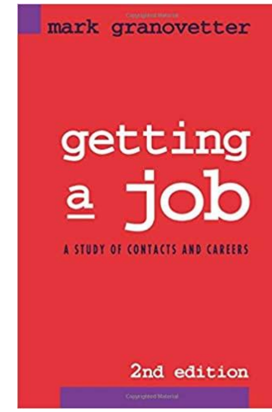


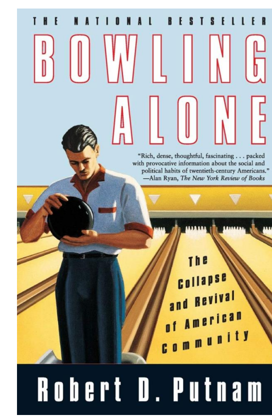
Connections Matter



This Jan. 26, 1965, file photo shows Mildred Loving and her husband Richard P Loving. Bernard S. Cohen, who successfully challenged a Virginia law banning interracial marriage.

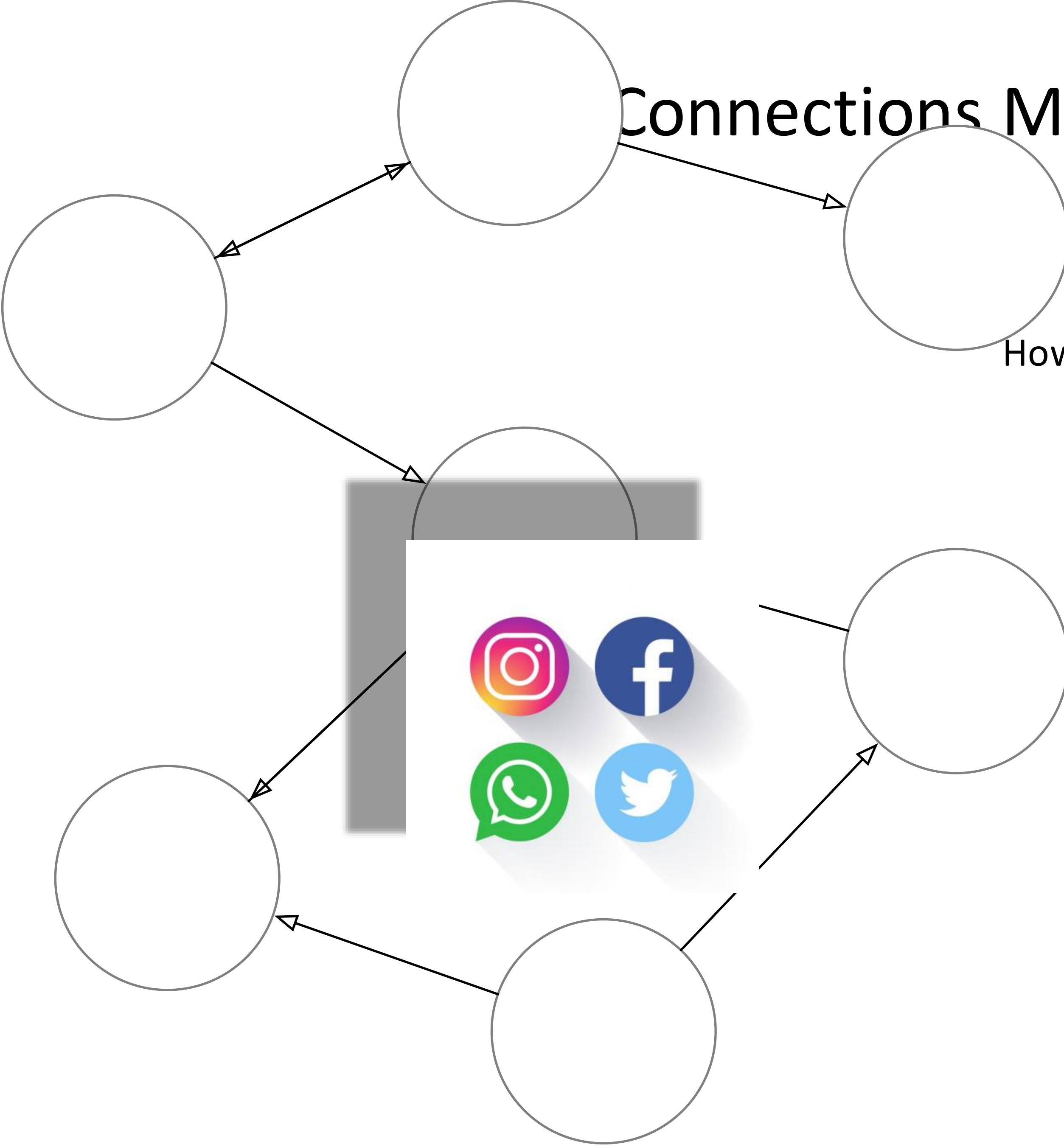






Connections Matter

How we connect is changing



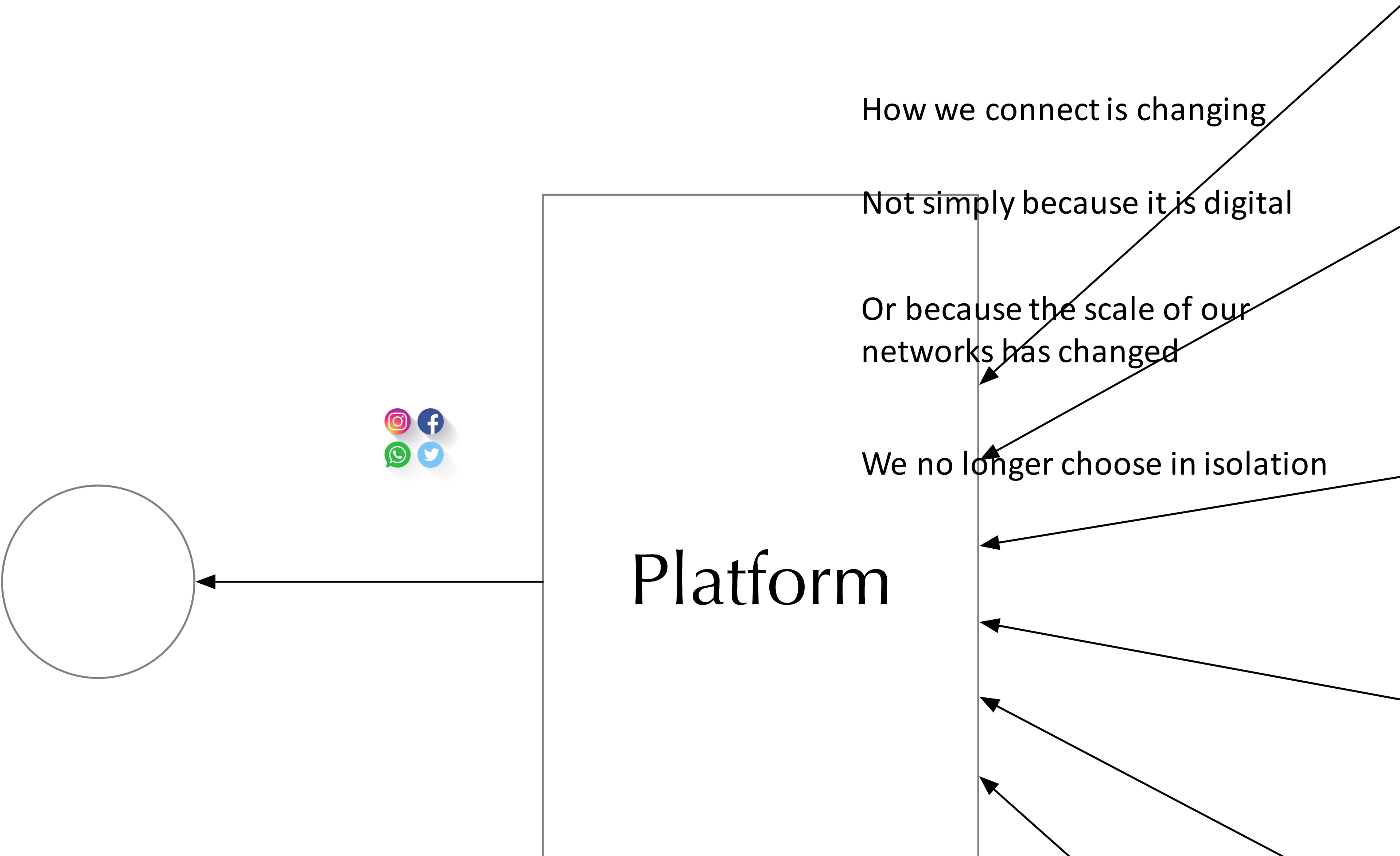
Social Connections Matter

How we connect is changing

Not simply because it is digital

Or because the scale of our networks has changed

We no longer choose in isolation



Social Connections Matter

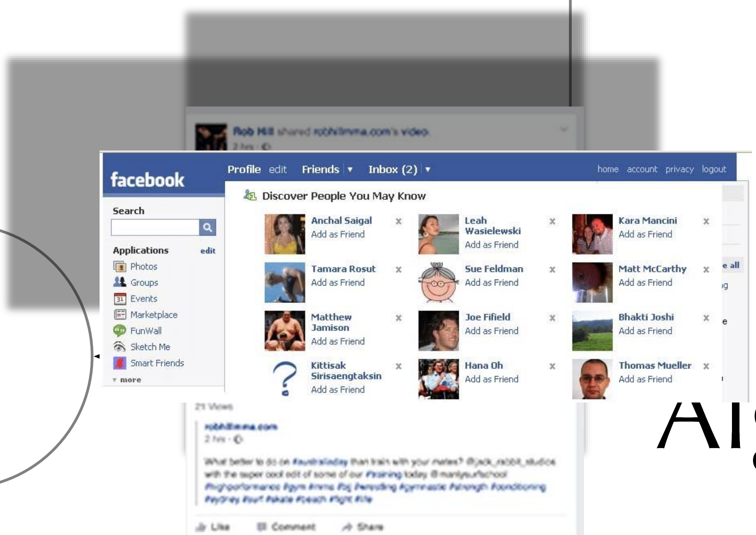
How we connect is changing

mically influenced human

s

Which friend do I engage with?

Newsfeed



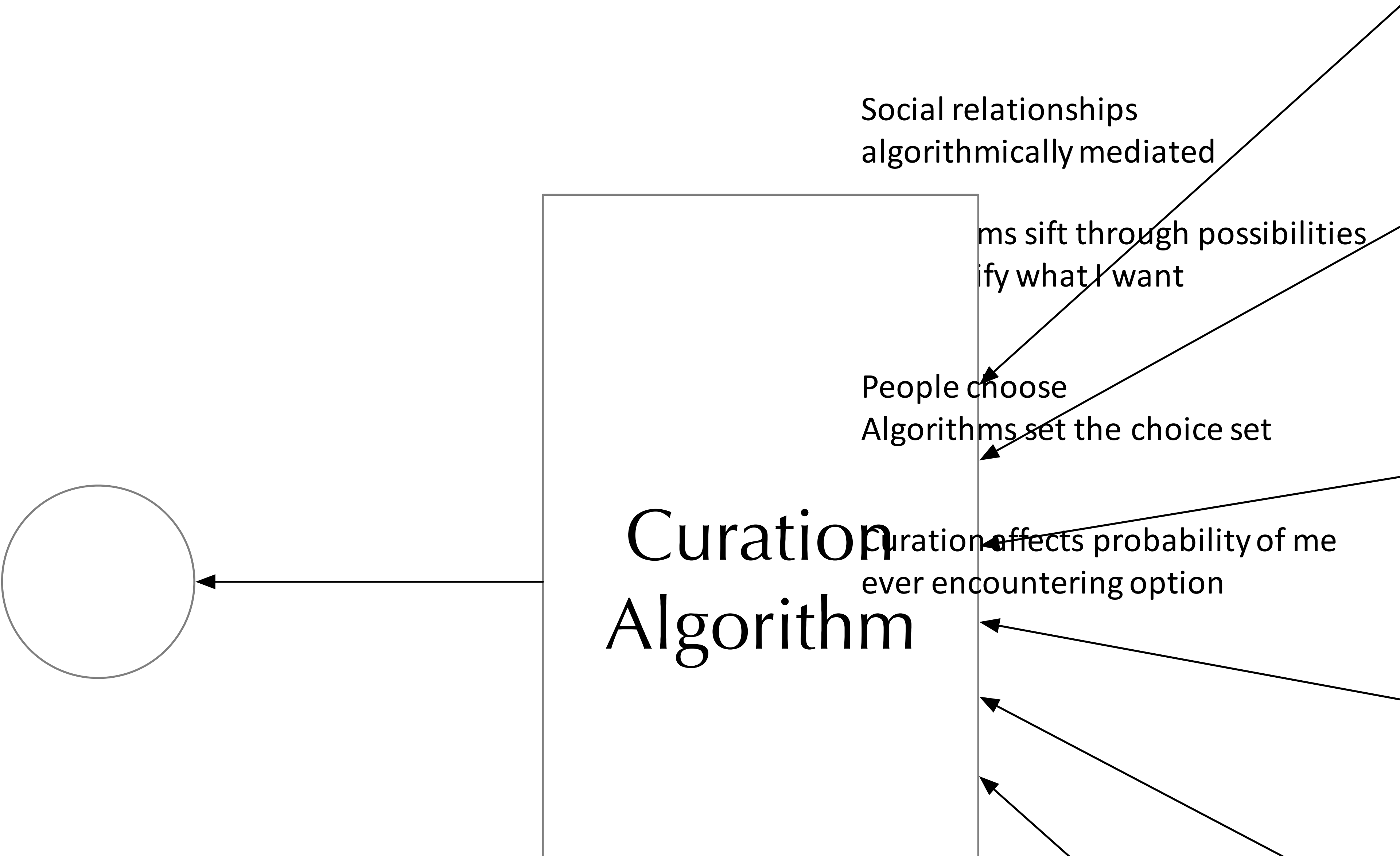
uration

Algorithm

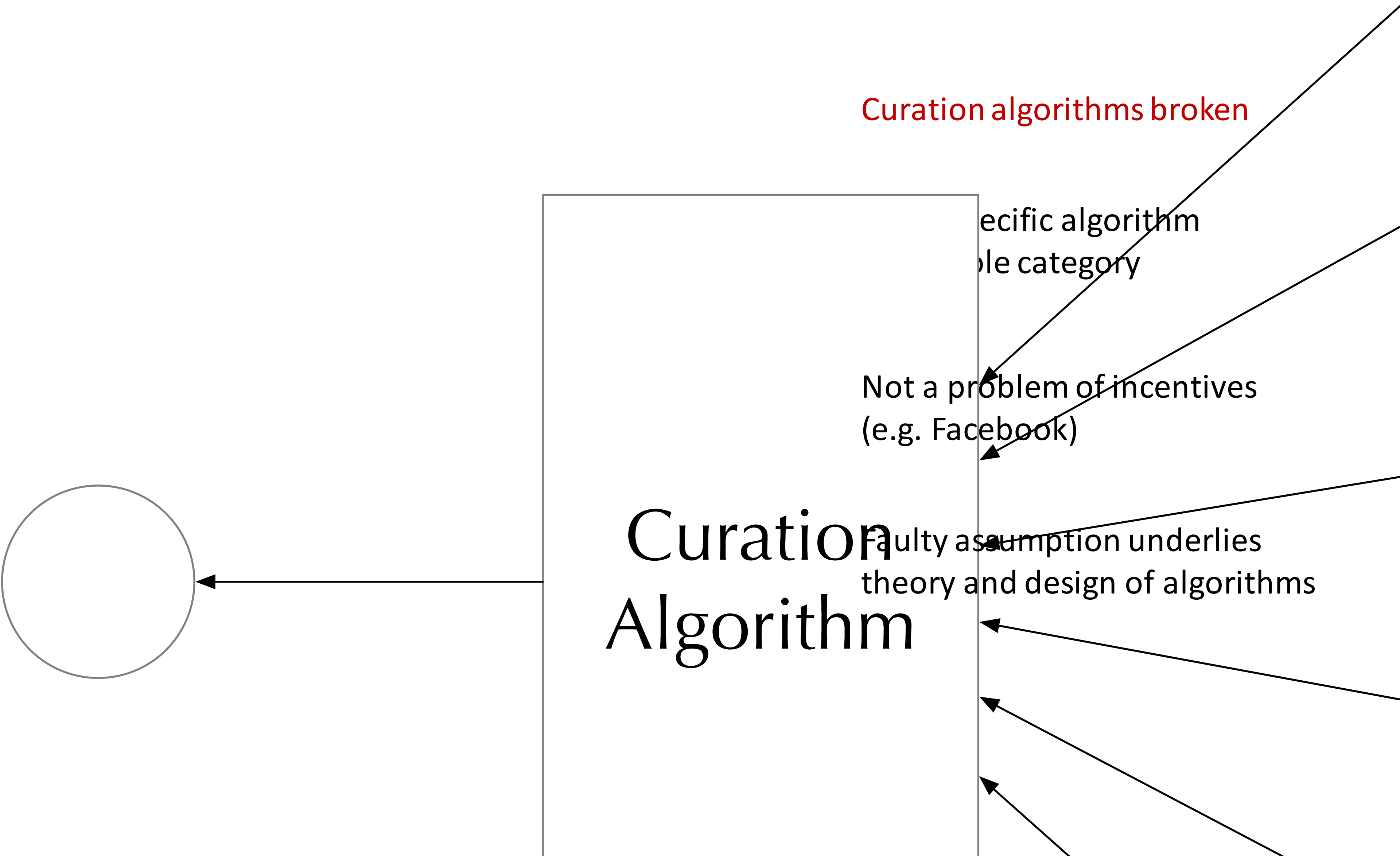
Who do I become friends with?

PYMK

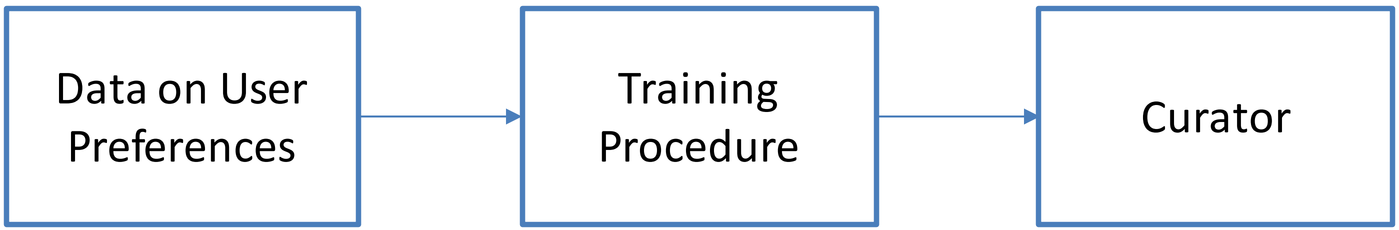
Curation Algorithms



Argument Today



Curation Algorithms



Important known problems
- Training data non-random
- Distribution shift
- Exploit/explore tradeoff

There is a more basic problem

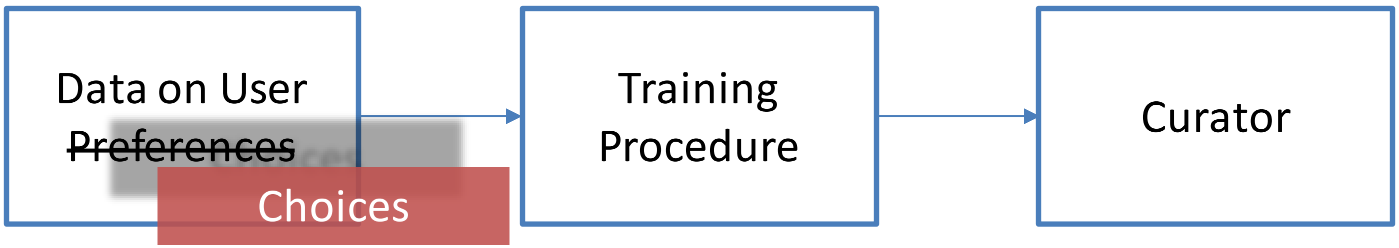
Curation Algorithms

Important known problems

- Training data non-random
- Distribution shift
- Exploit/explore tradeoff

There is a more basic problem

Curation Algorithms



Algorithms implicitly equate choice with preference

Revealed preference

Intuitive assumption: building block of axiomatic choice theory

Unfortunately it is wrong…

Curation Algorithms

Data on User ~~Preferences~~ Choices → Training Procedure → Curator

Preference

Revealed Preference Fails

Not just in behavior you regret
End of a long day spouse says "You forgot to take out the trash!"

One failure of revealed preference particularly noteworthy

Ex: Ward, Zanna and Cooper (1974)

White subjects interview White and Black “job applicants”

They asked people what they thought about the Black applicant.
Note the year – no taboos
Many people were happy to say race-speficic things (“I would never hire a black person”)

Not the important part of the study

They measured several <u>behaviors</u>

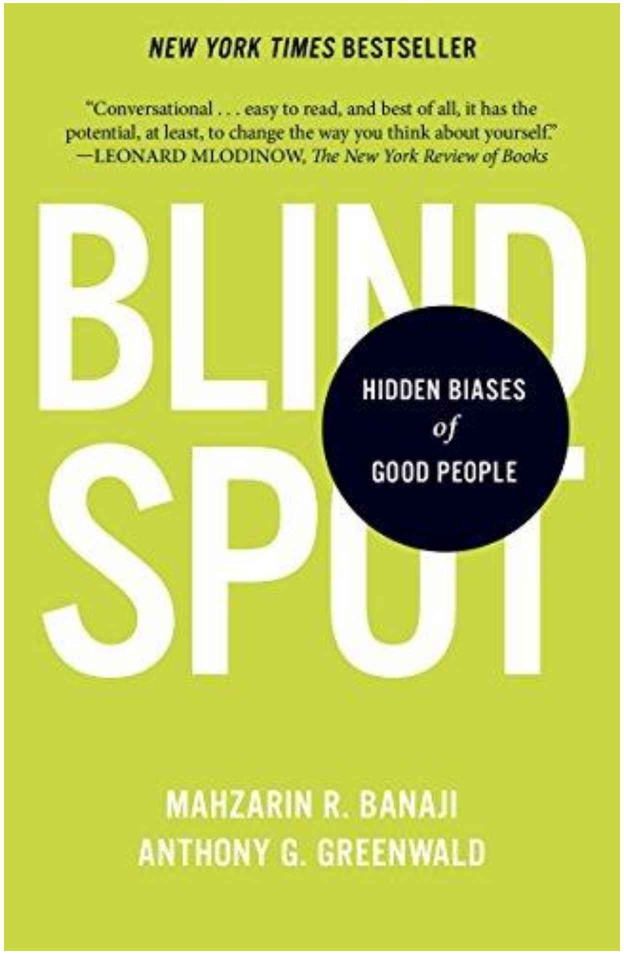
More racial bias in behaviors than attitudes

Speaking patterns

Interview length

Physical distance

Eye contact



Should sound familiar

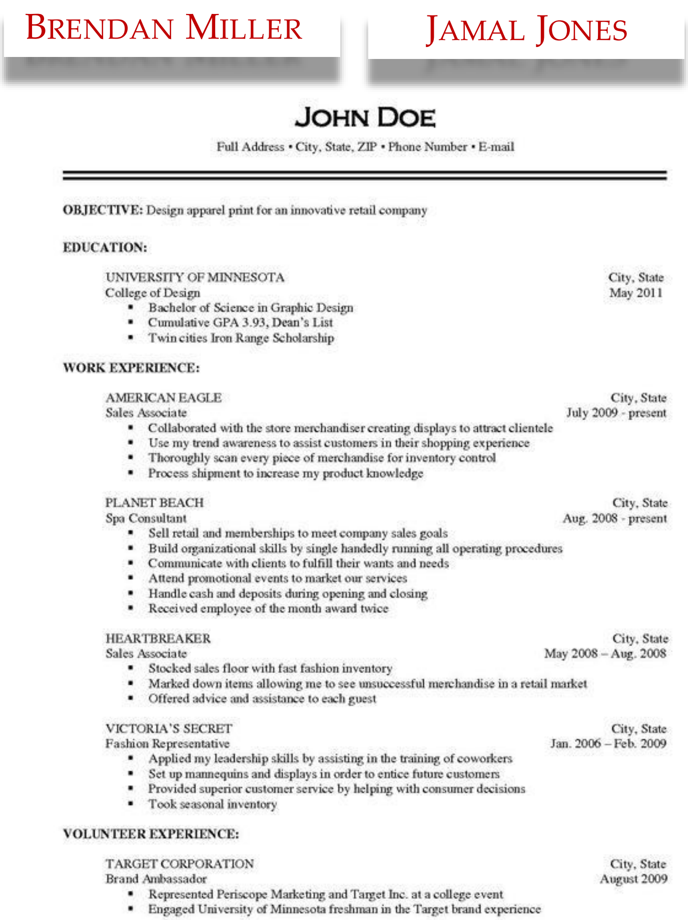
Implicit bias

People generally favor own-group

But behave more biasedly than preferences

Many of us are desire diversity and equality

Many of us inadvertently behave otherwise



BRENDAN MILLER

JAMAL JONES

JOHN DOE

Full Address • City, State, ZIP • Phone Number • E-mail

OBJECTIVE: Design apparel print for an innovative retail company

EDUCATION:

UNIVERSITY OF MINNESOTA — City, State
College of Design — May 2011
- Bachelor of Science in Graphic Design
- Cumulative GPA 3.93, Dean's List
- Twin cities Iron Range Scholarship

WORK EXPERIENCE:

AMERICAN EAGLE — City, State
Sales Associate — July 2009 - present
- Collaborated with the store merchandiser creating displays to attract clientele
- Use my trend awareness to assist customers in their shopping experience
- Thoroughly scan every piece of merchandise for inventory control
- Process shipment to increase my product knowledge

PLANET BEACH — City, State
Spa Consultant — Aug. 2008 - present
- Sell retail and memberships to meet company sales goals
- Build organizational skills by single handedly running all operating procedures
- Communicate with clients to fulfill their wants and needs
- Attend promotional events to market our services
- Handle cash and deposits during opening and closing
- Received employee of the month award twice

HEARTBREAKER — City, State
Sales Associate — May 2008 – Aug. 2008
- Stocked sales floor with fast fashion inventory
- Marked down items allowing me to see unsuccessful merchandise in a retail market
- Offered advice and assistance to each guest

VICTORIA'S SECRET — City, State
Fashion Representative — Jan. 2006 – Feb. 2009
- Applied my leadership skills by assisting in the training of coworkers
- Set up mannequins and displays in order to entice future customers
- Provided superior customer service by helping with consumer decisions
- Took seasonal inventory

VOLUNTEER EXPERIENCE:

TARGET CORPORATION — City, State
Brand Ambassador — August 2009
- Represented Periscope Marketing and Target Inc. at a college event
- Engaged University of Minnesota freshman in the Target brand experience

Should sound familiar

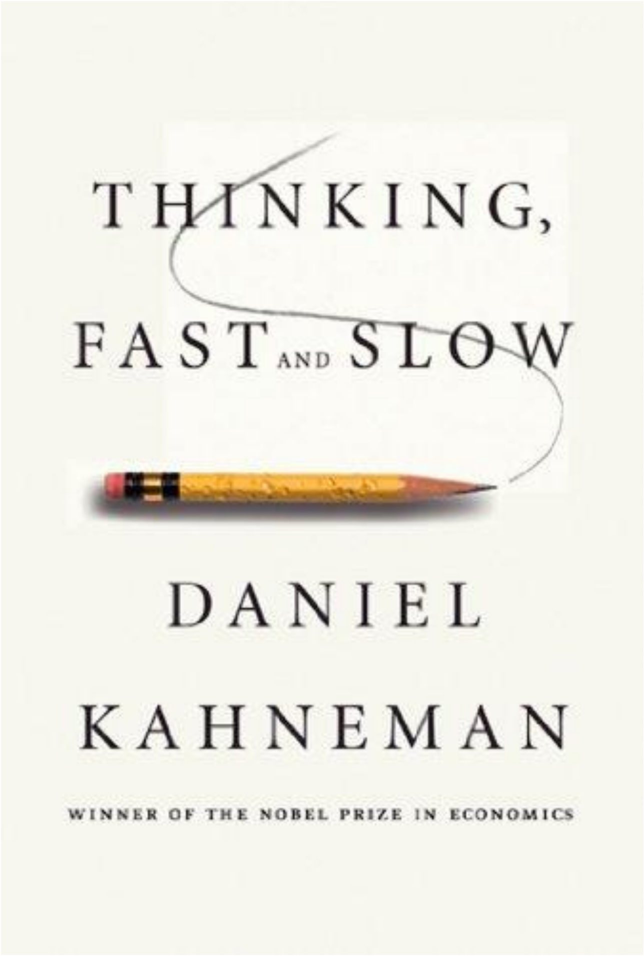
Implicit bias

People generally favor own-group

But behave more biasedly than preferences

Many of us are desire diversity and equality

Many of us inadvertently behave otherwise



Implicit bias has structure

It happens when we behave automatically

Low deliberation

Quick choices

Low consequences

Sound familiar?

Curation Algorithms

Trained <u>mostly</u> (but not always) in high automaticity contexts

…learn our outgroup favoritism from our behaviors

…which has more bias than our actual preferences

I click less, linger less on an outgroup friend’s post

Algorithm infers I ”like” it less

Curation Algorithms

Worse than mirroring bias

Outgroup posts get lower ranked and thus less likely to ever be seen

So posts are now <u>doubly</u> penalized

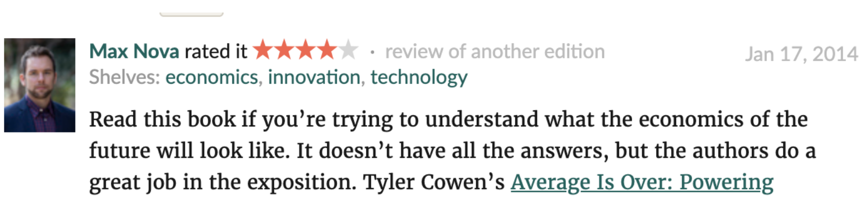
Human bias: Less likely to click when seen

Algorithmic bias: Less likely to be seen

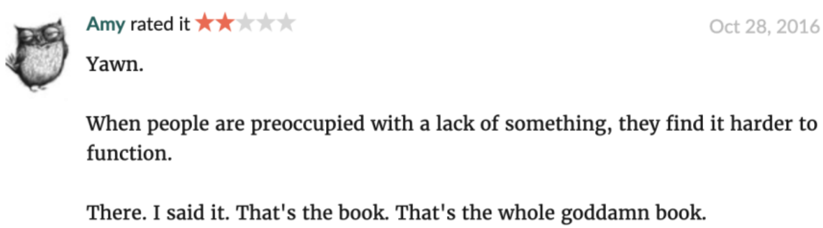
Importantly not universal: larger in contexts of automatic behavior

Lab Experiment

goodreads


Max Nova rated it ★★★★☆ · review of another edition Jan 17, 2014
Shelves: economics, innovation, technology

Read this book if you're trying to understand what the economics of the future will look like. It doesn't have all the answers, but the authors do a great job in the exposition. Tyler Cowen's Average Is Over: Powering


Amy rated it ★★☆☆☆ Oct 28, 2016

Yawn.

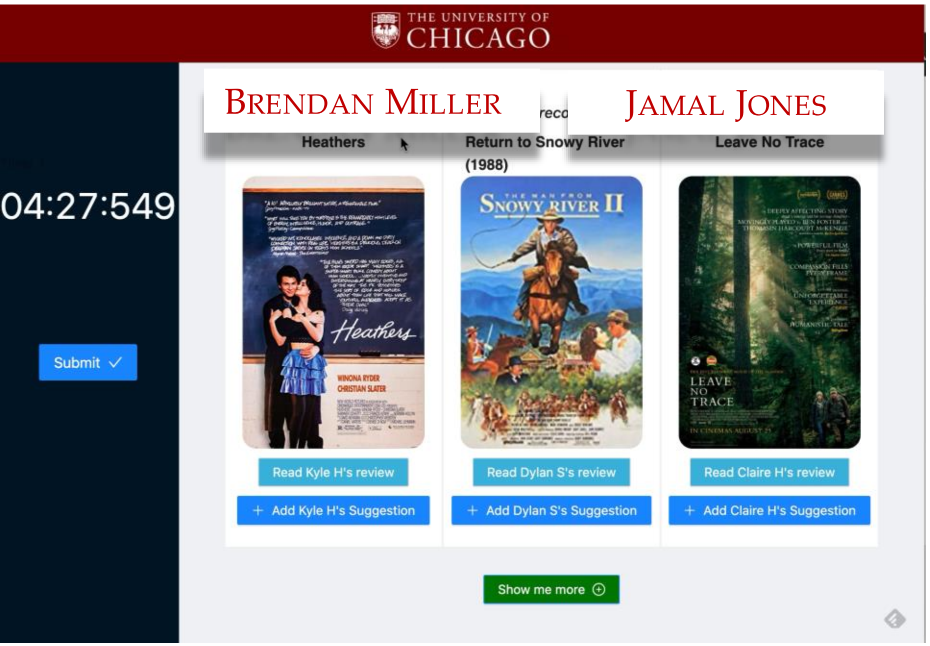
When people are preoccupied with a lack of something, they find it harder to function.

There. I said it. That's the book. That's the whole goddamn book.

Subjects engage with content generated by people

Books were a bit too close to home so we chose movies

Lab Experiment



Incentivized choices
So one of the movies they'd actually get free

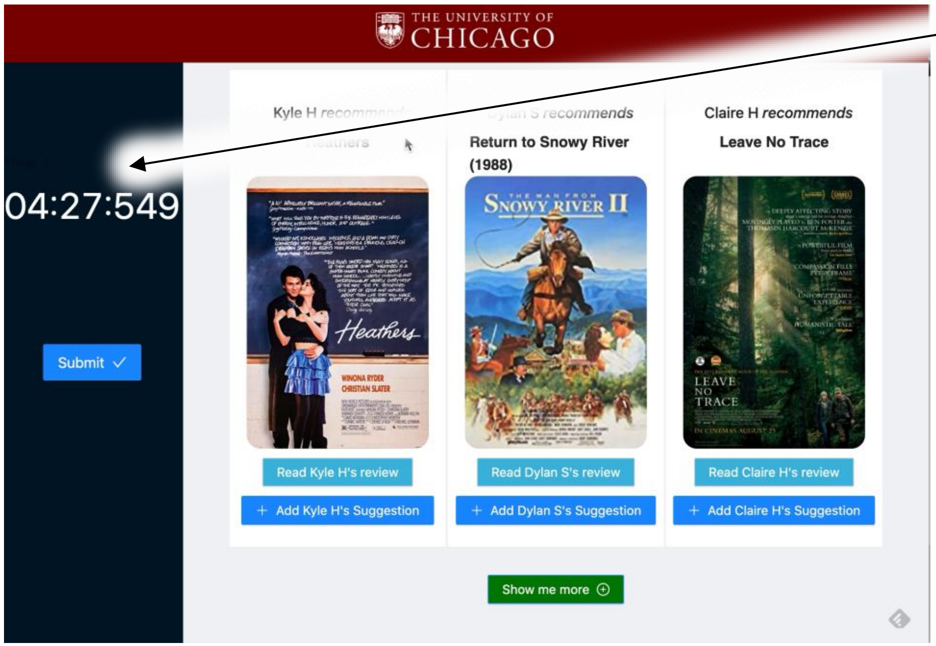
Random assignment of names meant we kept constant the true "likingness"

Question: does click rate depend on ingroup identity?

But recall we have one more <u>key</u> prediction

Bias depends on automaticity!

Lab Experiment



Automaticity manipulated through time pressure

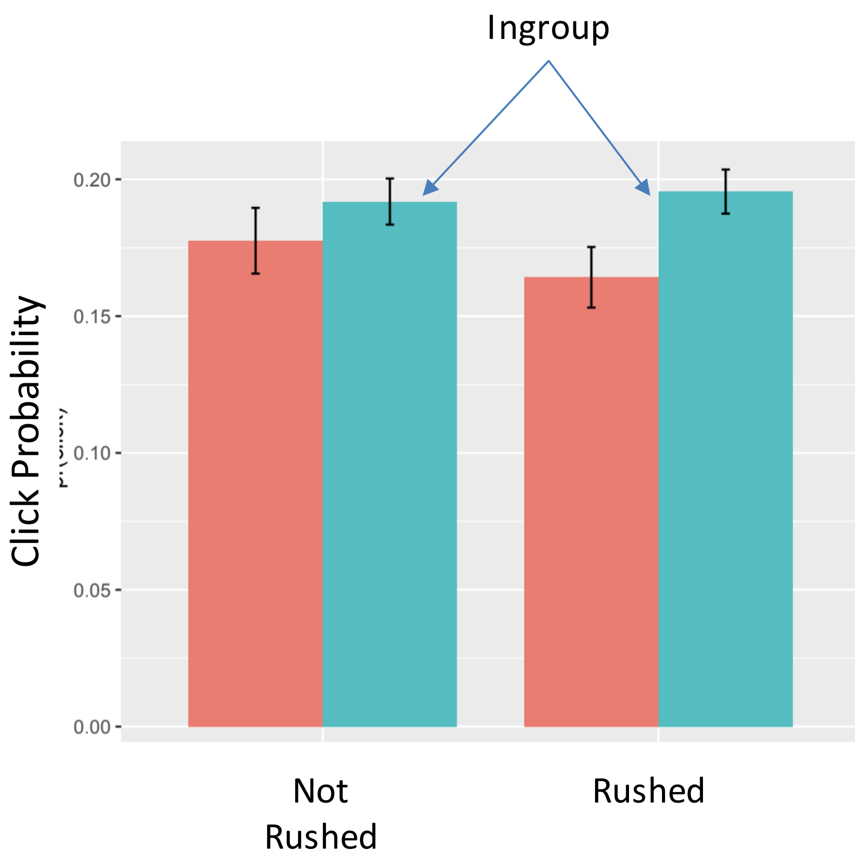
And count down clock

1000 subjects

Question: more clicks of ingroup recs?

Question: greater gap when rushed?

Click Rate differences



Here in-group is defined as same race or gender

Recall: Random assignment of movies meant we kept constant the true "likingness"

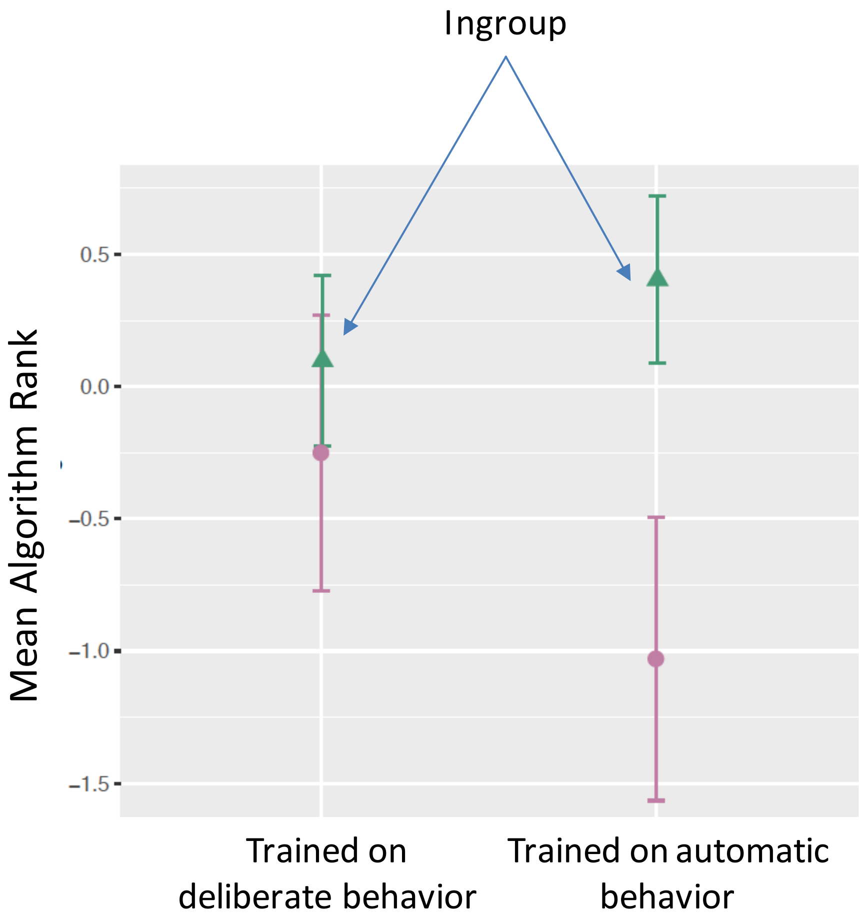
But of course all we have done is recreate a known psychological bias

Let's imagine these are two different worlds

The people are the same…so statistically their preferences are the same

Algorithm trained from one world's data
Or the other?

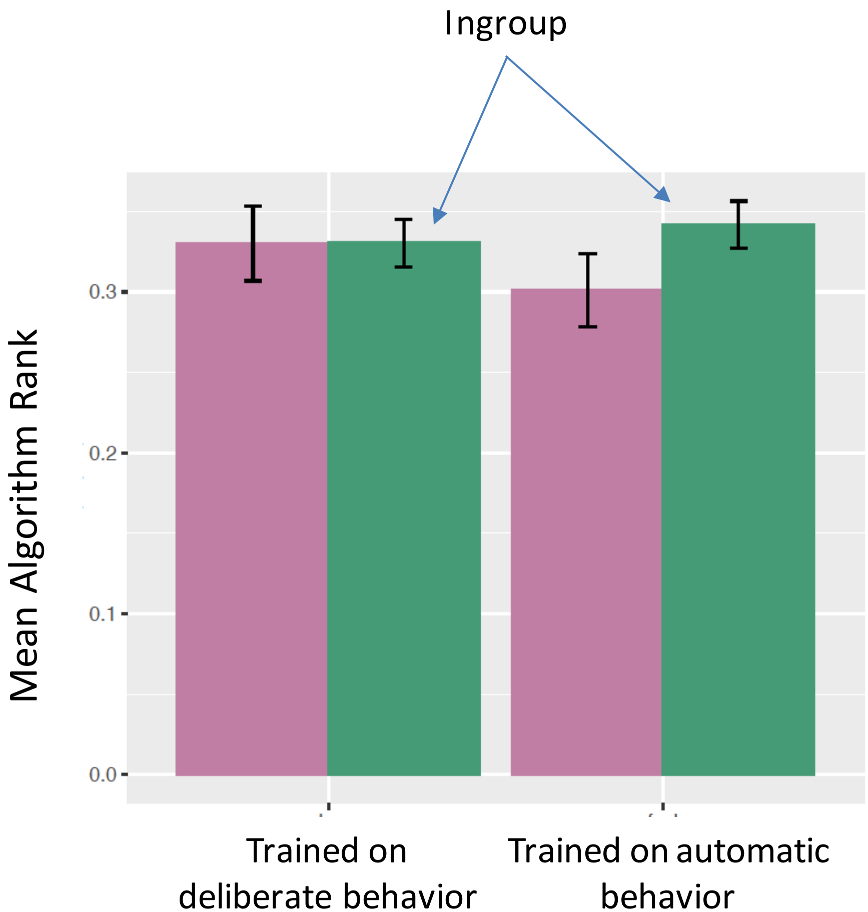
Algorithm Ranking



Outgroup posts ranked lower

But only if data comes from the rushed condition

Algorithm Ranking

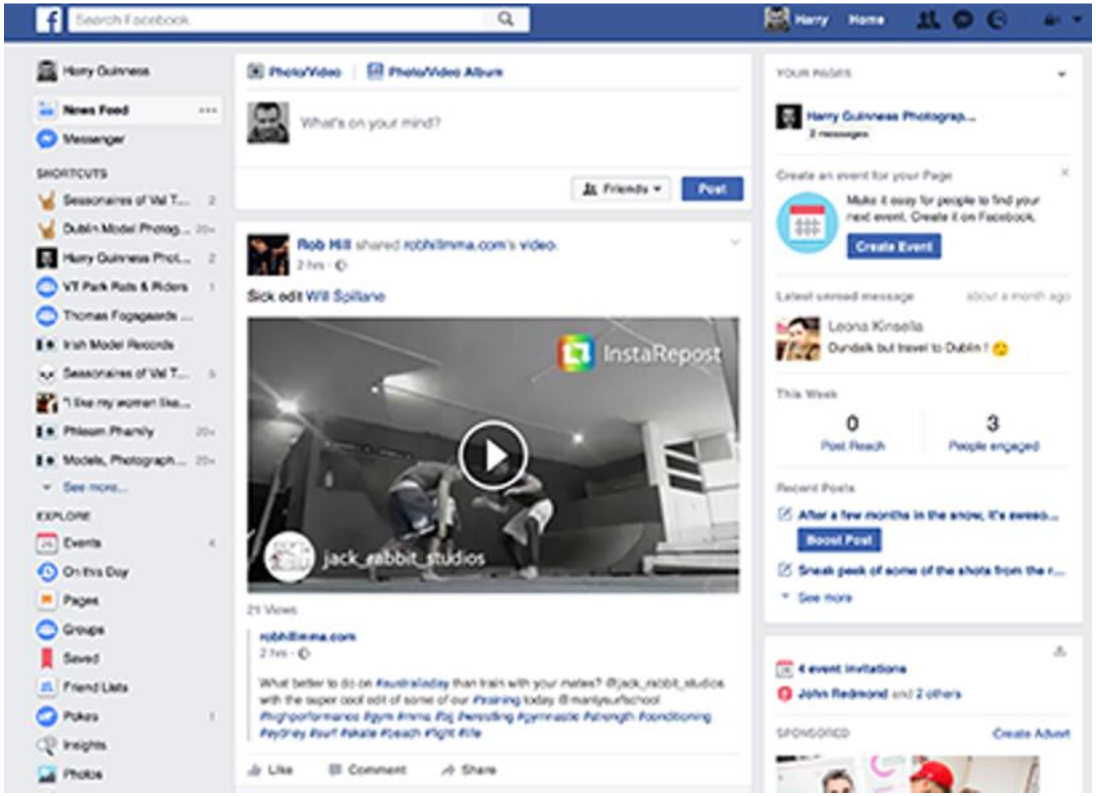


Outgroup posts ranked lower

But only if data comes from the rushed condition

Algorithmic bias has crept in here

But does it… in the world?



Recruit participants to screenshare while scrolling through FB

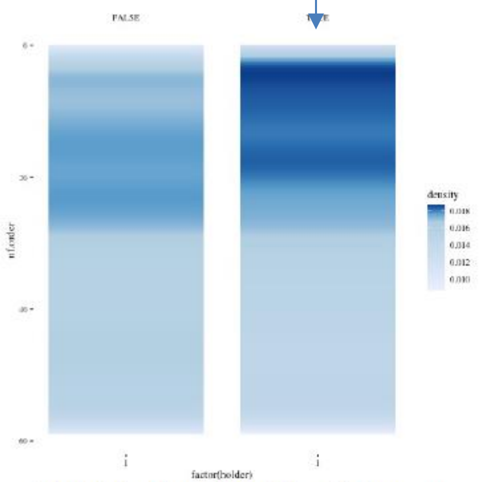
Track first 60 non-sponsored posts.

(Unbeknownst to participant) record race of friend behind each post

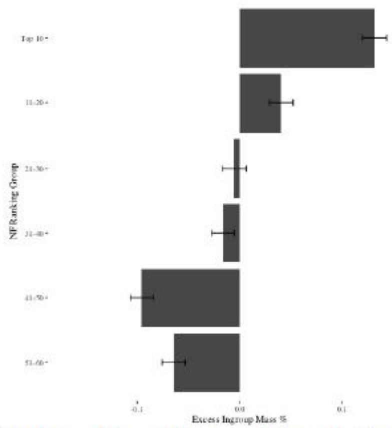
Look at how ingroup and outgroup posts rank

Ingroup and Outgroup post rankings

Ingroup



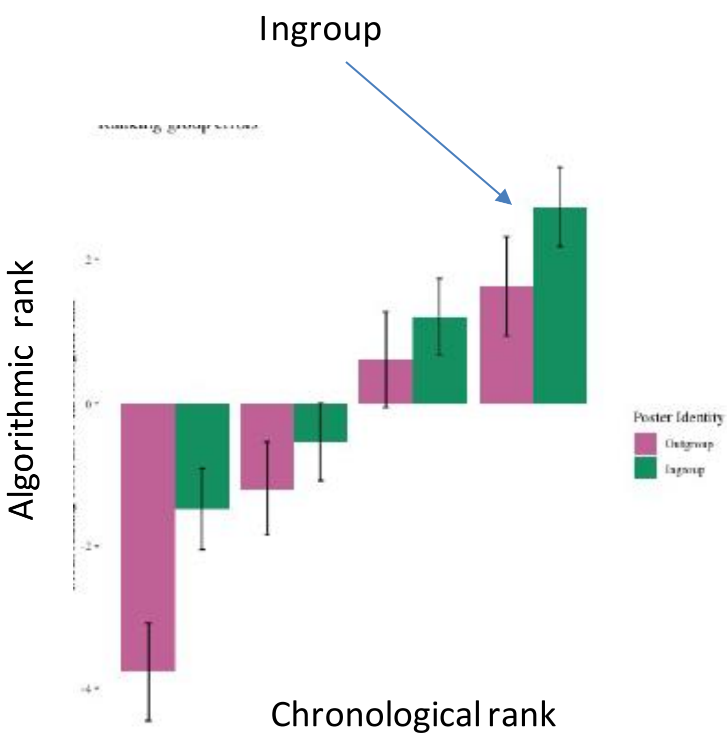
(a) Relative Intensity of Posts by in-group Status and Algorithmic Ranking



(b) Excess Mass of in-group Posts by Newsfeed Algorithmic Ranking

Ingroup posts ranked higher

Ingroup and Outgroup post rankings

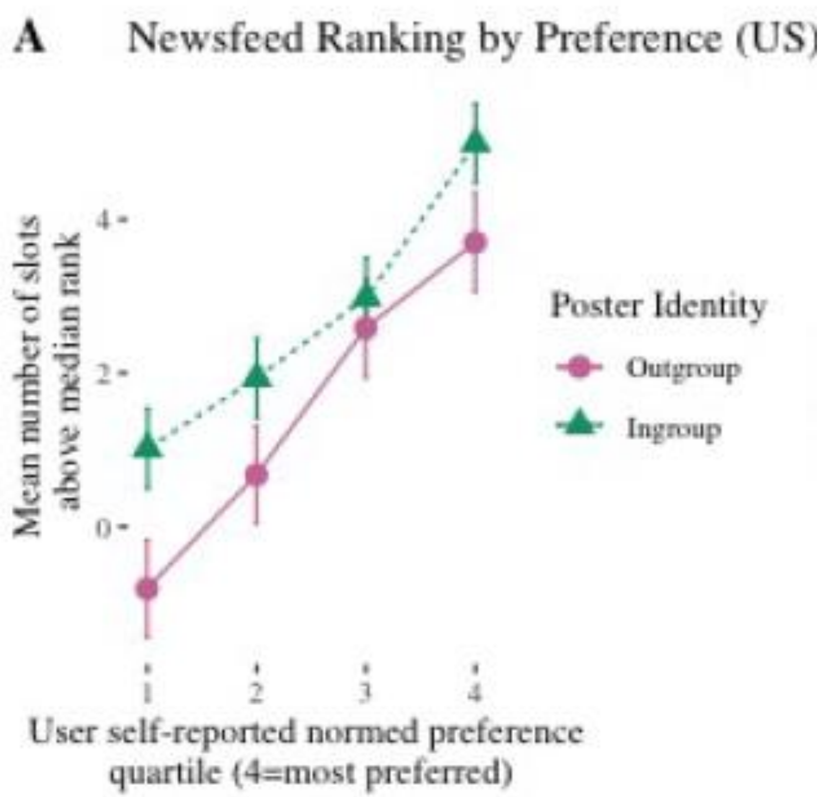


Ingroup posts ranked higher

Maybe actual preferences (not just algorithm) also have this bias?

Asked subjects
"how much do you want to see this post?" [1-7]

Ingroup and Outgroup post rankings

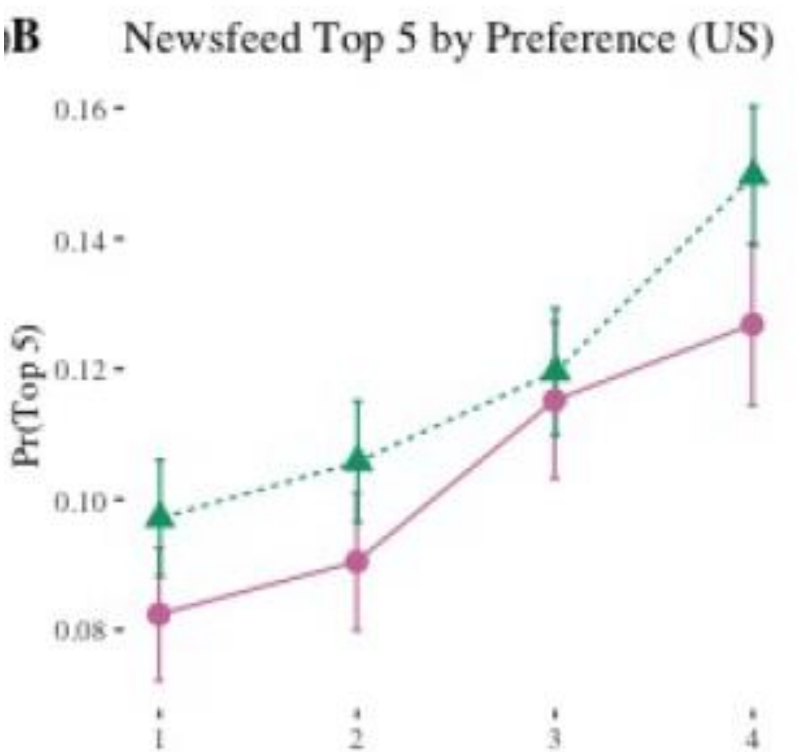


Ingroup posts ranked higher

Maybe actual preferences (not just algorithm) also have this bias?

Asked subjects
“how much do you want to see this post?” [1-7]

Ingroup and Outgroup post rankings

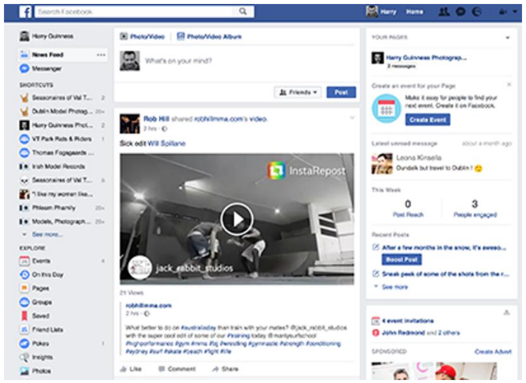


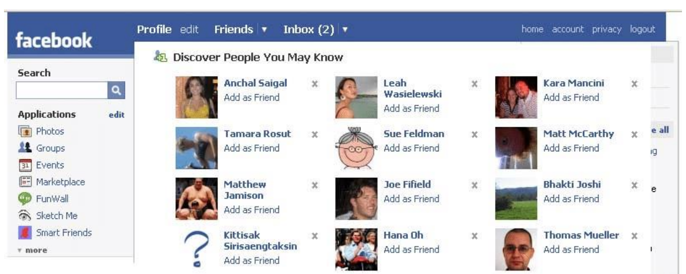
Ingroup posts ranked higher

Ingroup posts much higher probability of being in the top 5

Is this happening because of automaticity?

How do we unpack?

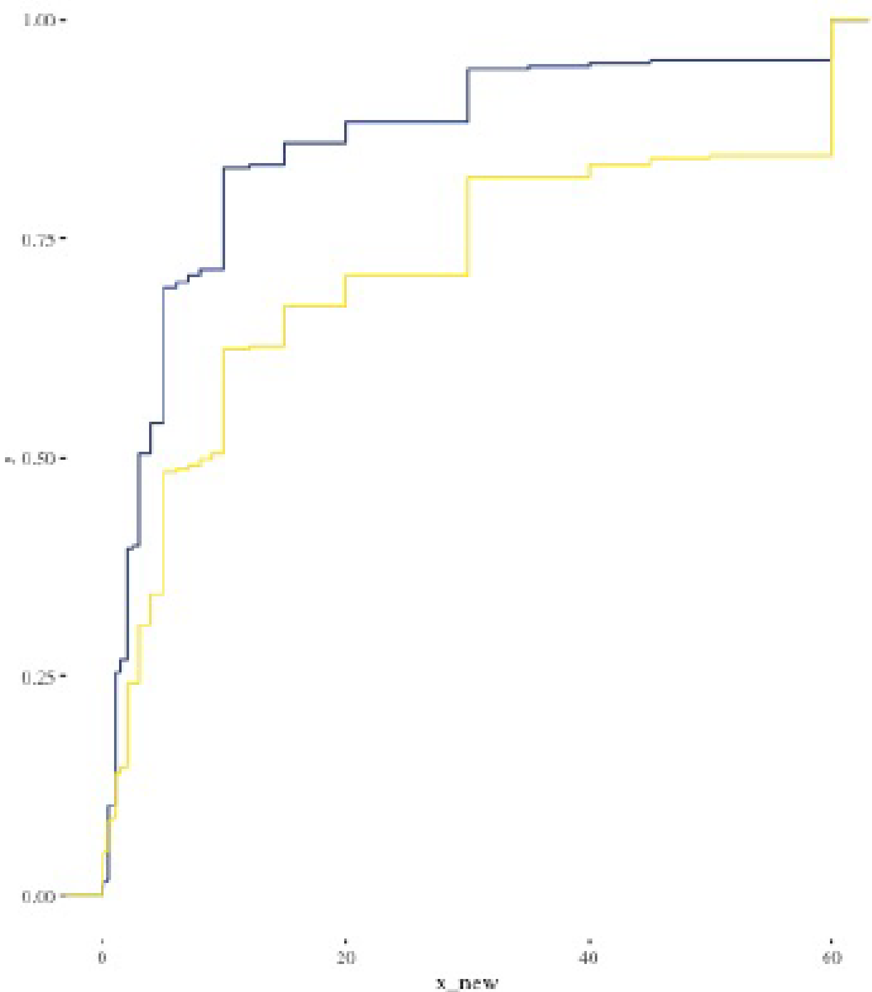




Difference between PYMK and Newsfeed

Newsfeed involves very quick decisions

PYMK involves very deliberate decisions



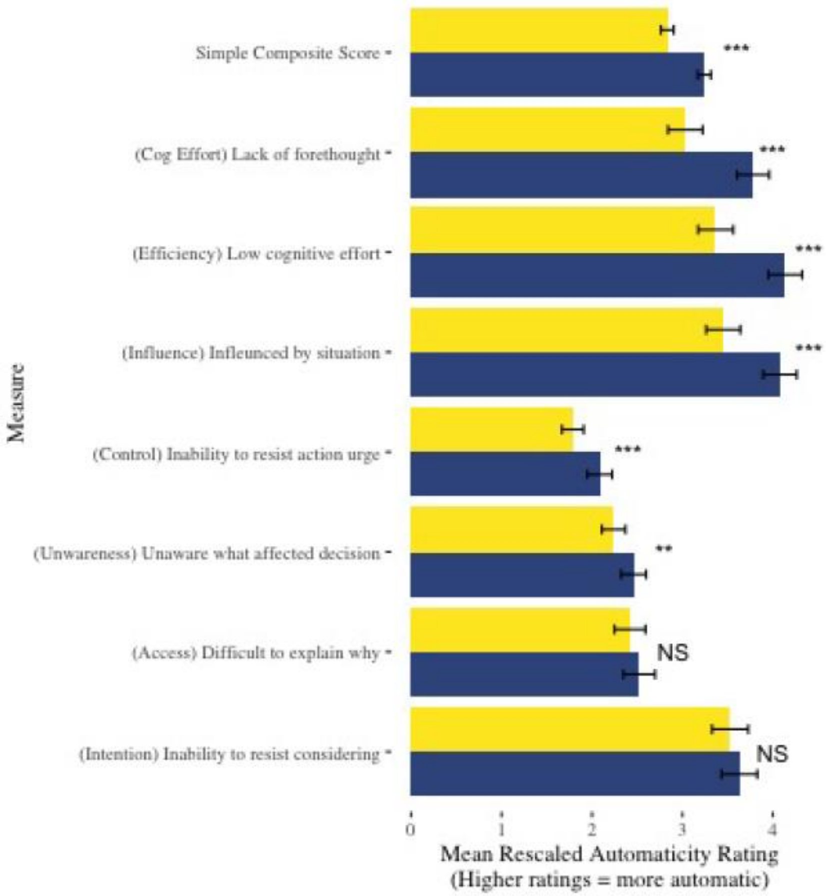
(d) Speed (time to decide in seconds) CDF

Difference between PYMK and Newsfeed

Newsfeed involves very quick decisions

PYMK involves very deliberate decisions

Similar results on many other measures of deliberateness



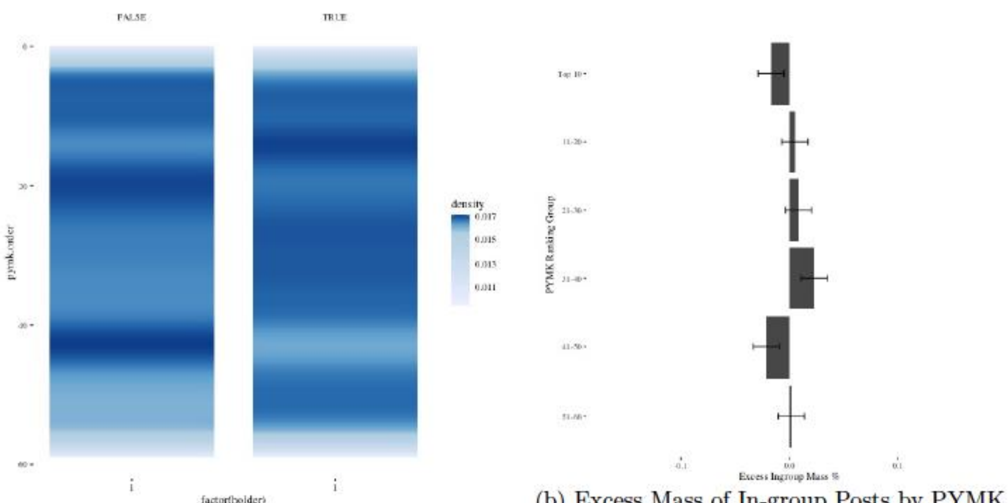
Difference between PYMK and Newsfeed

Newsfeed involves very quick decisions

PYMK involves very deliberate decisions

Similar results on many other measures of deliberateness

Figure 5: Relationship between PYMK Algorithmic Ranking and in-group Status of Posts



(a) Relative Intensity of Posts by In-group status and Algorithmic Ranking

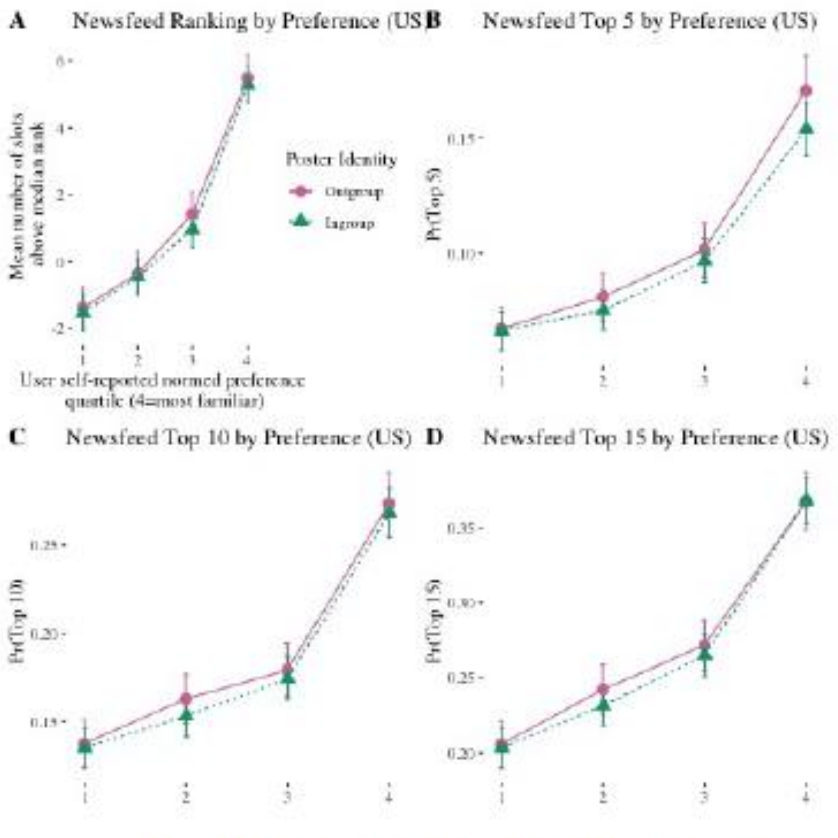
(b) Excess Mass of In-group Posts by PYMK Algorithmic Ranking

Audit PYMK. Compare where suggestion ranks to….

First 60 recommendations, “how familiar are you with this person?” [1-7]

(Unbeknownst to participant) record race of recommended friend

<u>No ingroup bias</u>



(a) Ranking Boost Above Median Rank

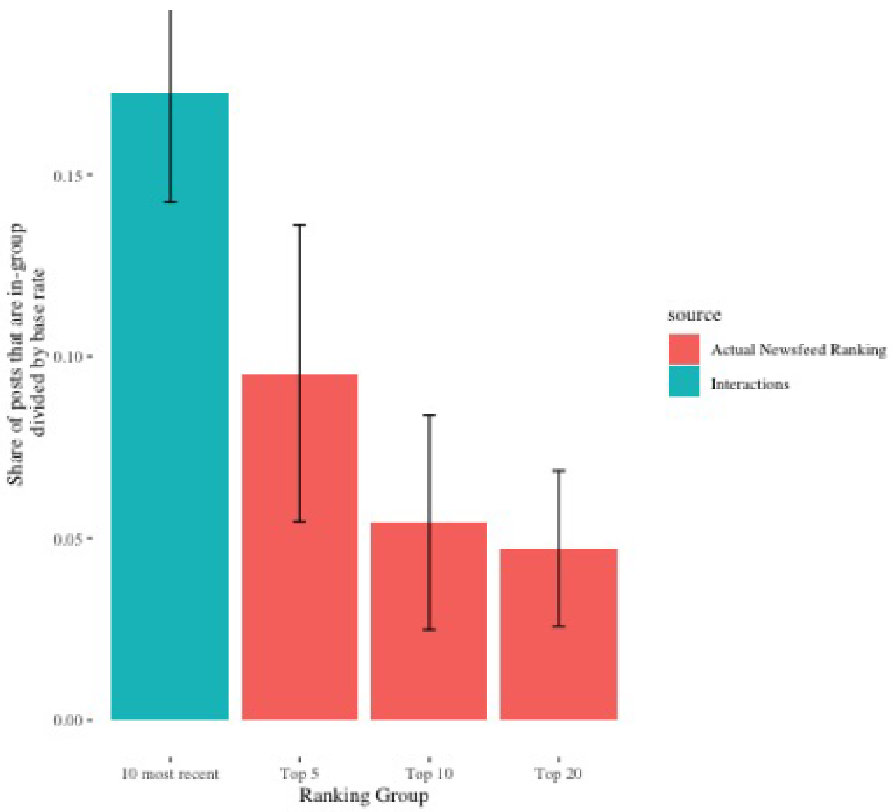
Audit PYMK. Compare where suggestion ranks to….

First 60 recommendations, "how familiar are you with this person?" [1-7]

(Unbeknownst to participant) record race of recommended friend

No ingroup bias

Figure 8: Share Recent Interactions In-Group versus Newsfeed Posts



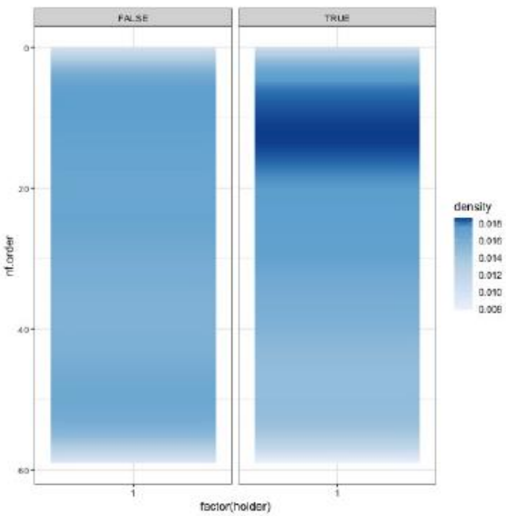
Returning to Newsfeed we have suggestive evidence of algorithm's increasing bias

We see outgroup posts are ranked lower which means…

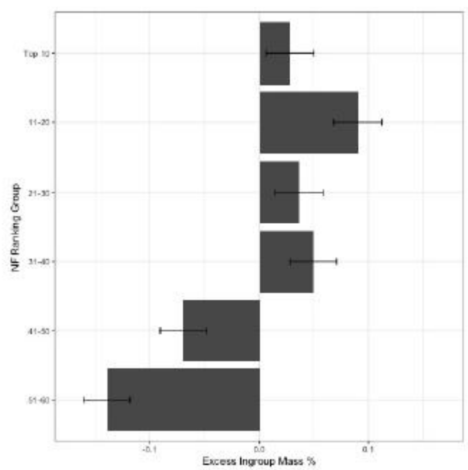
1) They are less seen
2) But also less liked when seen

So we looked at 10 most recent likes/interactions a person had

Figure D.10: NF in-group Posts Higher



(a) Two Single Heat Maps by in-group



(b) Excess Mass by Ranking

Finally we ran all of these in a different context….

India: Ingroup defined by religion. Hindu and Muslim

Summary

- In controlled lab conditions…
  - Automatic behavior produces bias
  - Algorithms trained on that data recreate bias
- Meaningful problem on Facebook
  - Large ingroup bias in newsfeed
  - But not on PYMK where behavior is less automatic
  - Repeats on an audit study of Muslim/Hindus in India
- Opens up important questions of..
  - Algorithm design
  - Social impact