



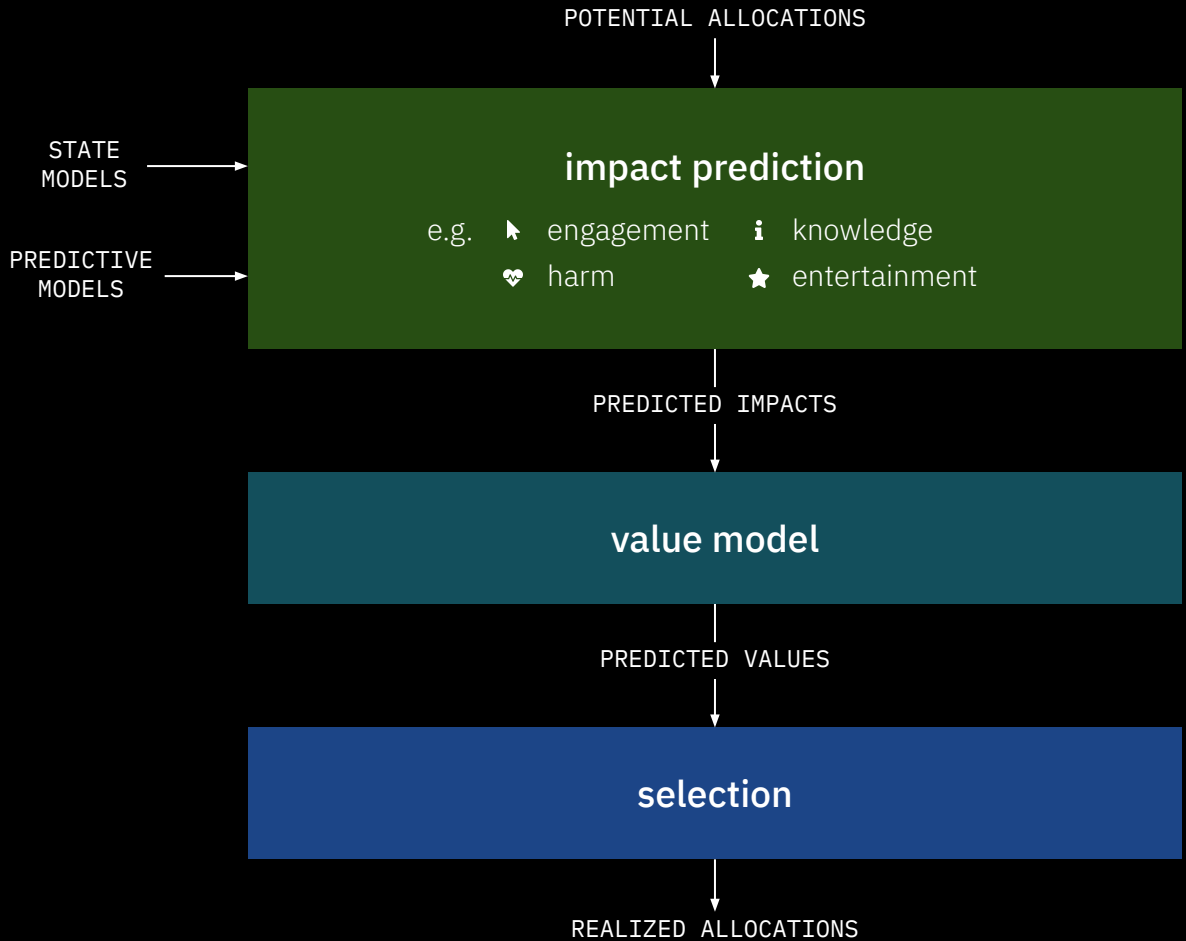
# Bridging-Based Ranking

Luke Thorburn

September 2023

**bridging**

An increase in mutual understanding and trust across divides, creating space for productive conflict, deliberation, or cooperation.



$$\begin{aligned} \text{value} = & 1 \cdot \text{Pr}(\textit{like}) + 3 \cdot \text{Pr}(\textit{comment}) + 10 \cdot \text{Pr}(\textit{share}) \\ & + 4 \cdot (\textit{expected dwell time}) - 6 \cdot \text{Pr}(\textit{fake} \\ & \textit{account}) \\ & - 7 \cdot \text{Pr}(\textit{low quality news}) - 10 \cdot \text{Pr}(\textit{clickbait}) \end{aligned}$$

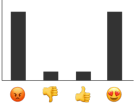
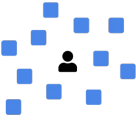
**Motifs**

**Surveys**

**Content**

**Formal Measures**

# Motifs

	diverse approval	response bimodality	exposure diversity ⚠️
“dividing”			
“bridging”			

# Content

# Surveys

# Formal Measures

# Motifs

	diverse approval	response bimodality	exposure diversity ⚠️
“dividing”			
“bridging”			

# Content

# Surveys



How does **Carmel Richardson's** tweet make you feel about people or groups on the Left?

Much worse -2      Worse -1      The same as before 0      Better 1      Much better 2





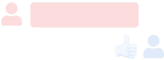
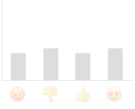

---

How does **Carmel Richardson's** tweet make you feel about people or groups on the Right?

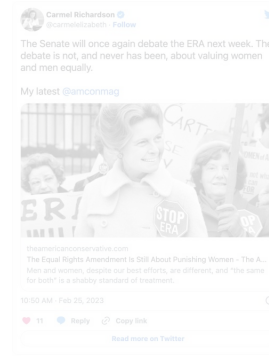
Much worse -2      Worse -1      The same as before 0      Better 1      Much better 2

# Formal Measures

# Motifs

	diverse approval	response bimodality	exposure diversity 
“dividing”			
“bridging”			

# Surveys



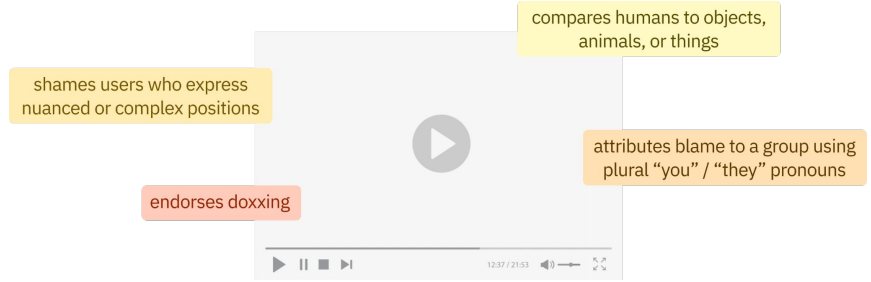
How does **Carmel Richardson's** tweet make you feel about people or groups on the Left?

Much worse -2    Worse -1    The same as before 0    Better 1    Much better 2

How does **Carmel Richardson's** tweet make you feel about people or groups on the Right?

Much worse -2    Worse -1    The same as before 0    Better 1    Much better 2

# Content



compares humans to objects, animals, or things

shames users who express nuanced or complex positions

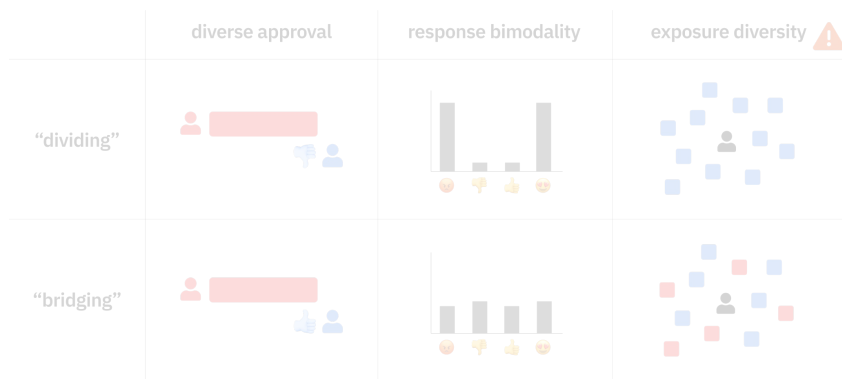
attributes blame to a group using plural “you” / “they” pronouns

endorses doxxing

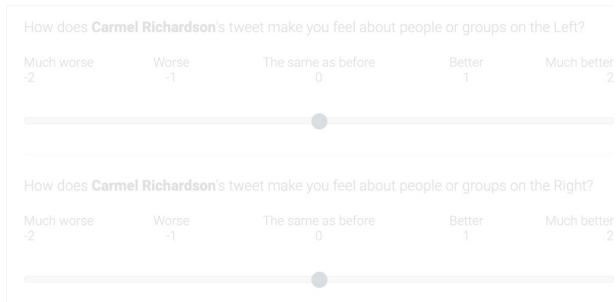
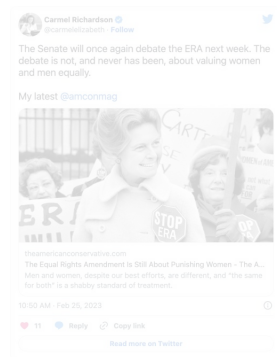
# Formal Measures



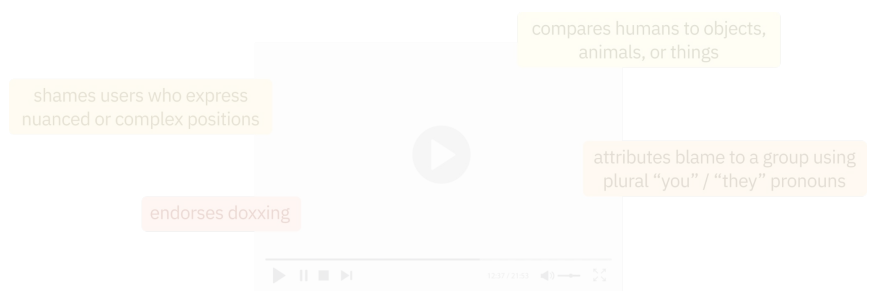
# Motifs



# Surveys



# Content



shames users who express nuanced or complex positions

compares humans to objects, animals, or things

endorses doxxing

attributes blame to a group using plural “you” / “they” pronouns

# Formal Measures



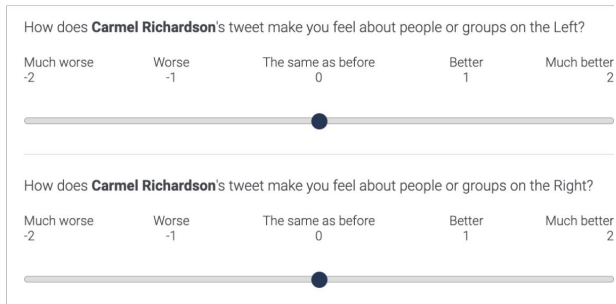
$$\frac{\mathbf{P}(R(\text{start}) = v \mid R(\text{end}) \in G_i^+)}{\sum_{j=1}^M \mathbf{P}(R(\text{start}) = v \mid R(\text{end}) \in G_j^+)} \rightarrow 0.34$$

A formal metric for quantifying conflict in a relation model, such as *random walk controversy*.

# Motifs

	diverse approval	response bimodality	exposure diversity ⚠️
“dividing”			
“bridging”			

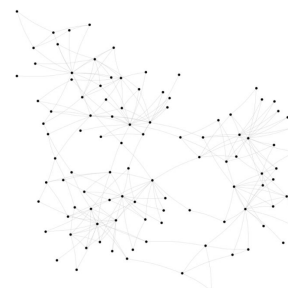
# Surveys



# Content

- shames users who express nuanced or complex positions
- compares humans to objects, animals, or things
- endorses doxxing
- attributes blame to a group using plural “you” / “they” pronouns

# Formal Measures



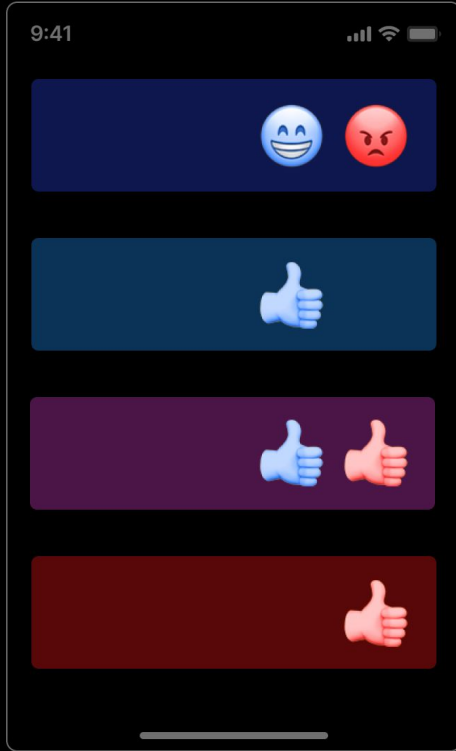
$$\frac{\mathbf{P}(R(\text{start}) = v \mid R(\text{end}) \in G_i^+)}{\sum_{j=1}^M \mathbf{P}(R(\text{start}) = v \mid R(\text{end}) \in G_j^+)} \rightarrow 0.34$$

A formal metric for quantifying conflict in a relation model, such as *random walk controversy*.

**diverse approval**

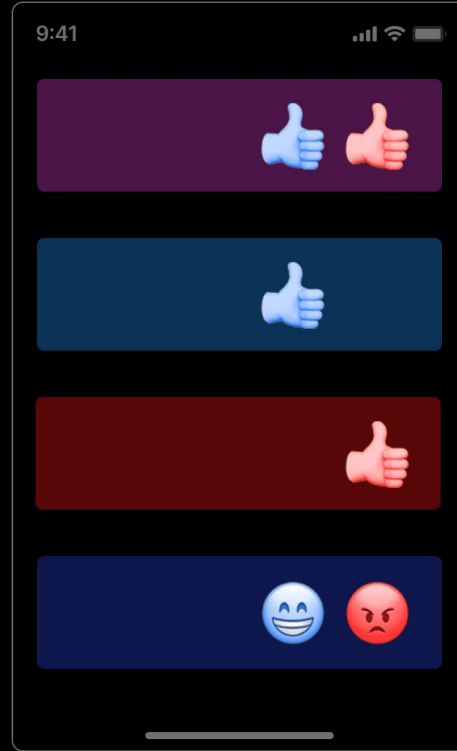
A positive response from a diverse group of people.

**ENGAGEMENT**  
based ranking  
for [blue bar]



**BRIDGING**  
based ranking  
for [blue bar]

*"diverse approval"*



Why?

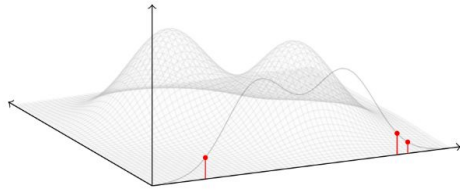
How?

Does it work?

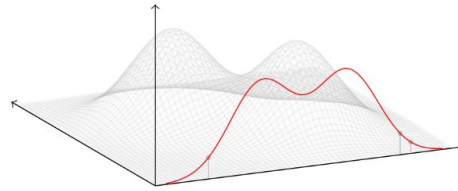
Why?

How?

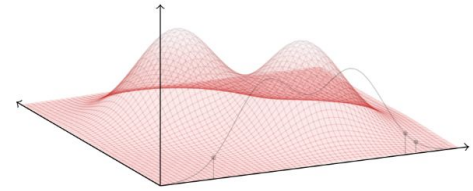
Does it work?



(a) "extremism"

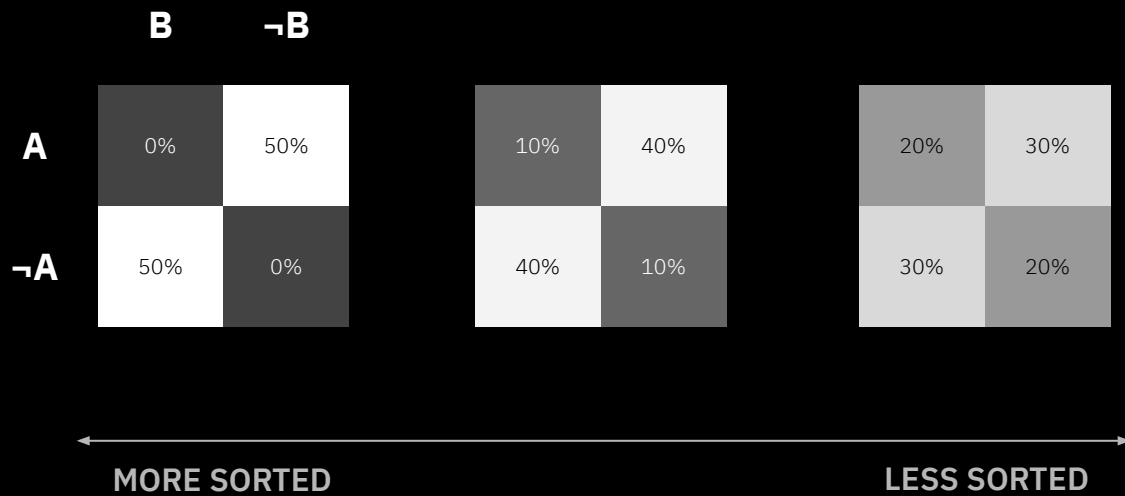


(b) "polarization"

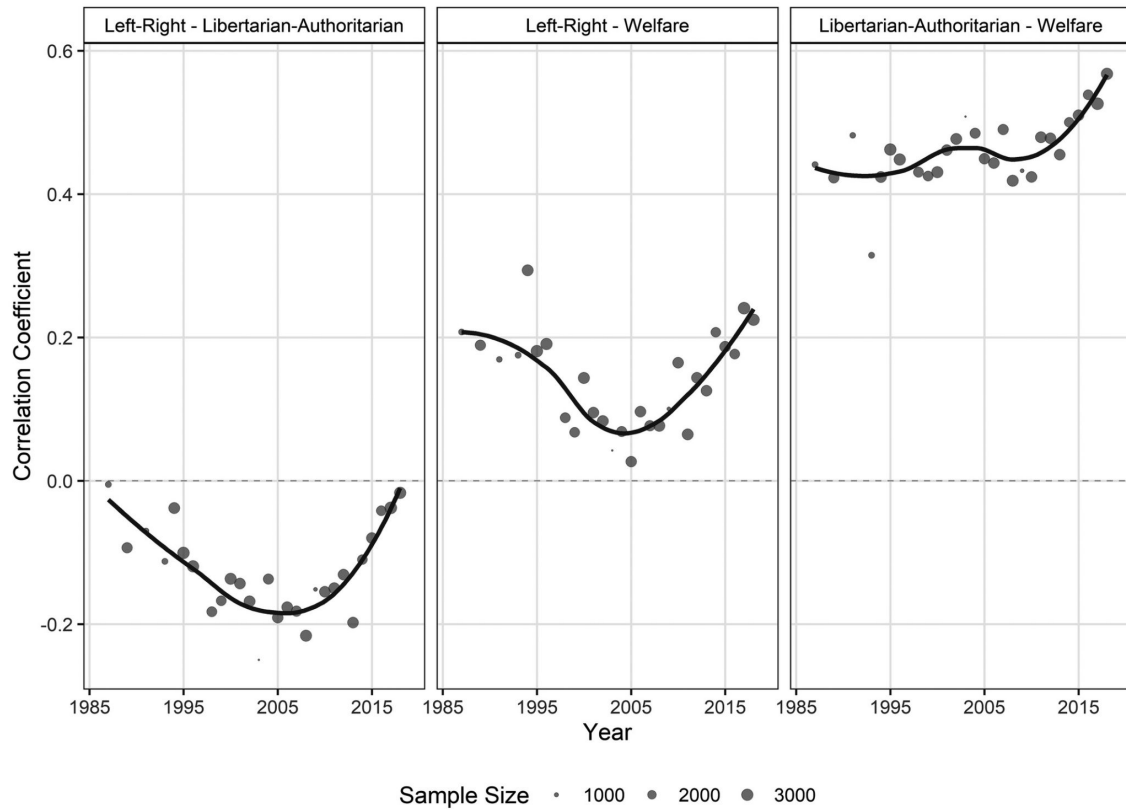


(c) "sorting"

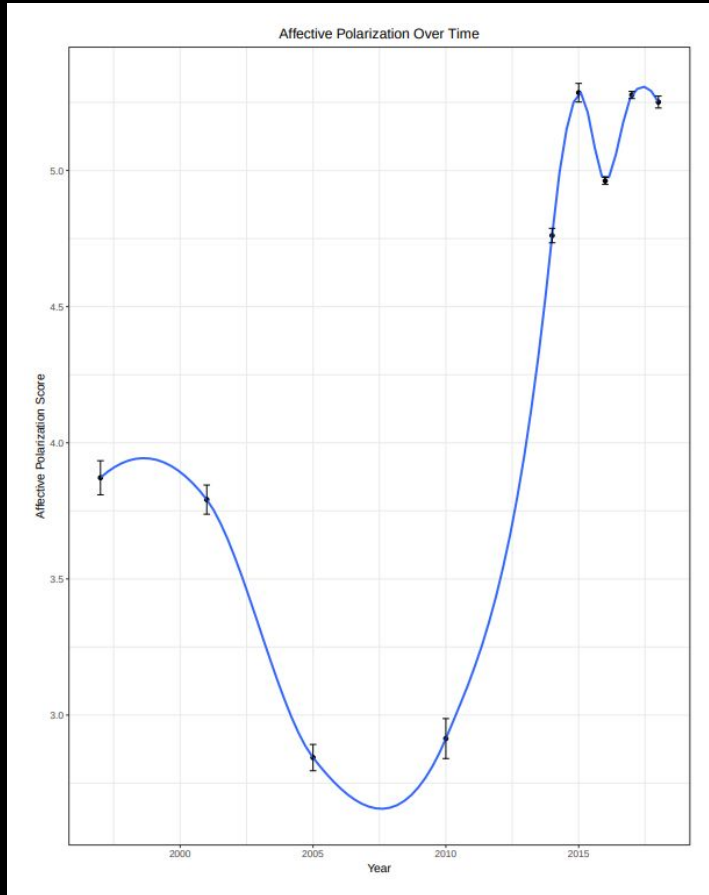
Luke Thorburn, Maria Polukarov and Carmine Ventre, "Probabilistic Foundations of Partisan (Un)Sorting" (forthcoming).







Stuart Perrett, "A divided kingdom? Variation in polarization, sorting, and dimensional alignment among the British public, 1986–2018" (2021).



Nahema Marchal and David Watson, "The Rise of Partisan Affective Polarization in the British Public" (2019).

## CONFLICT STUDIES, COLLECTIVE INTELLIGENCE

pigeonholing  
civil war onset  
“surprising validation”  
“complicate the narratives”  
perception gaps  
humanisation / contact theory  
common ground / symmarchy  
correlation discounting  
consilience / wisdom of the crowd

## POLITICAL THEORY, DEMOCRATIC THEORY

**safeguards against factionalism**

James Madison, *The Federalist No. 10* (1787)

**“intersection of social circles”**

Georg Simmel, *Soziologie* (1908)

**“overlapping memberships”**

Robert Dahl, *A Preface to Democratic Theory* (1956)

**“overlapping consensus”**

John Rawls, *Political Liberalism* (1993)

**“diverse diversities”**

Amartya Sen, *Identity and Violence* (2006)

## CONFLICT STUDIES, COLLECTIVE INTELLIGENCE

pigeonholing  
civil war onset  
“surprising validation”  
“complicate the narratives”  
perception gaps  
humanisation / contact theory  
common ground / symmarchy  
correlation discounting  
consilience / wisdom of the crowd

## POLITICAL THEORY, DEMOCRATIC THEORY

safeguards against factionalism  
Instead of...  
James Madison, *The Federalist No. 10* (1787)  
“how predictable and uninformative that  
someone like that would think something so  
evil and foolish”  
Georg Simmel, *Sozialogie* (1908)  
“overlapping consensus”  
Robert Dahl, *A Preface to Democratic Theory* (1956)  
...we respond:  
“overlapping consensus”  
John Rawls, *Political Liberalism* (1993)  
“if someone like that disagrees with me,  
maybe I had better rethink”  
Amartya Sen, *Identity and Violence* (2006)

→ **Breaking Up the Echo**, Cass Sunstein (2012)

→ **Filtering for Serendipity — Extremism, “Filter Bubbles” and “Surprising Validators”**, Richard Reisman (2012)

## CONFLICT STUDIES, COLLECTIVE INTELLIGENCE

pigeonholing  
civil war onset  
“surprising validation”  
“**complicate the narratives**”  
perception gaps  
humanisation / contact theory  
common ground / symmarchy  
correlation discounting  
consilience / wisdom of the crowd

## POLITICAL THEORY, DEMOCRATIC THEORY

**safeguards against factionalism**

James Madison, *The Federalist No. 10* (1787)

**“intersection of social circles”**

“When people encounter complexity, they become more curious and less closed off to new information.”

Robert Dahl, *A Preface to Democratic Theory* (1956)

Recommends:

**“amplifying contradictions”**

John Rawls, *Political Liberalism* (1993)

**“diverse diversities”**

Amartya Sen, *Identity and Violence* (2006)

## CONFLICT STUDIES, COLLECTIVE INTELLIGENCE

pigeonholing  
civil war onset  
“surprising validation”  
“complicate the narratives”  
perception gaps  
humanisation / **contact theory**  
common ground / symmarchy  
correlation discounting  
consilience / wisdom of the crowd

A meta-analysis with **515 studies** and more than **250,000 subjects** demonstrates that intergroup contact typically reduces prejudice (**mean  $r = -.21$** ). Allport's original conditions for optimal contact – equal status, common goals, no intergroup competition, and authority sanction – facilitate the effect but are **not necessary conditions**. There are other positive outcomes of intergroup contact, such as **greater trust and forgiveness for past transgressions**. These contact effects occur **not only for ethnic groups but also for such other groups as homosexuals, the disabled and the mentally ill**. ... Moreover, these effects **typically generalize** beyond the immediate outgroup members in the situation to the whole outgroup, other situations, and even to other outgroups not involved in the contact. They also **appear to be universal** – across nations, genders, and age groups.

## CONFLICT STUDIES, COLLECTIVE INTELLIGENCE

pigeonholing  
civil war onset  
“surprising validation”  
“complicate the narratives”  
perception gaps  
humanisation / contact theory  
common ground / symmarchy  
correlation discounting  
consilience / wisdom of the crowd

## POLITICAL THEORY, DEMOCRATIC THEORY

**safeguards against factionalism**

James Madison, *The Federalist No. 10* (1787)

**“intersection of social circles”**

Georg Simmel, *Soziologie* (1908)

**“overlapping memberships”**

Robert Dahl, *A Preface to Democratic Theory* (1956)

**“overlapping consensus”**

John Rawls, *Political Liberalism* (1993)

**“diverse diversities”**

Amartya Sen, *Identity and Violence* (2006)

## CONFLICT STUDIES, COLLECTIVE INTELLIGENCE

pigeonholing  
civil war onset  
“surprising validation”  
“complicate the narratives”  
perception gaps  
humanisation / contact theory  
common ground / symmarchy  
correlation discounting  
consilience / wisdom of the crowd

## POLITICAL THEORY, DEMOCRATIC THEORY

**safeguards against factionalism**

James Madison, *The Federalist No. 10* (1787)

**“intersection of social circles”**

Georg Simmel, *Soziologie* (1908)

**“overlapping memberships”**

Robert Dahl, *A Preface to Democratic Theory* (1956)

**“overlapping consensus”**

John Rawls, *Political Liberalism* (1993)

**“diverse diversities”**

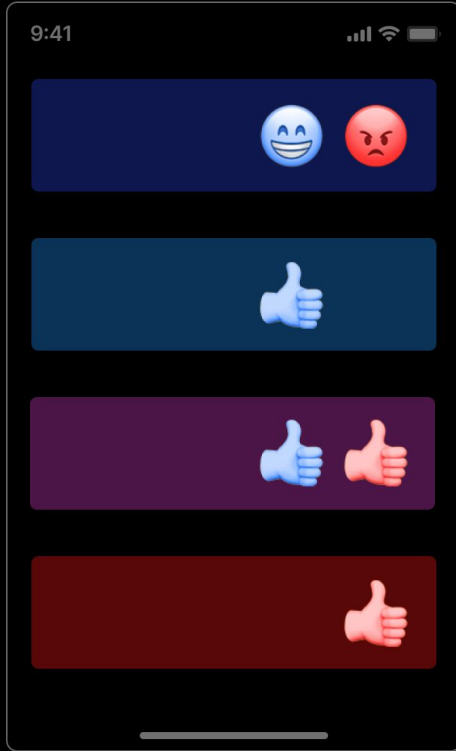
Amartya Sen, *Identity and Violence* (2006)

**Sorting “increases inter-party conflict and makes cross-party compromise more difficult”.**

Abrams and Fiorina (2015)

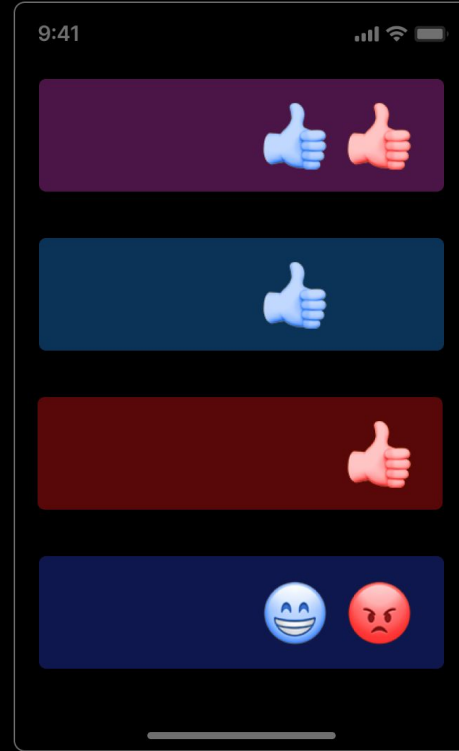


**ENGAGEMENT**  
based ranking  
for [blue bar]

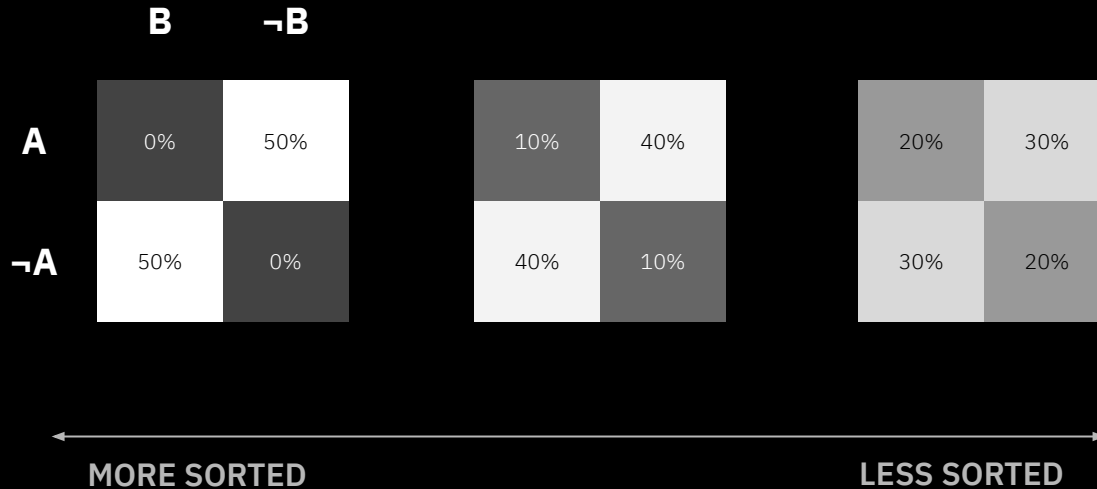


**BRIDGING**  
based ranking  
for [blue bar]

*"diverse approval"*



# diverse approval $\approx$ reduced sorting





## Misunderstandings about **bridging**

Bridging is **X** not eliminating conflict or promoting a particular viewpoint...

...it is about promoting valuable forms of conflict.

Bridging is **X** not simply showing people content from the other side...

...research shows that in some contexts this can make conflict worse.

## Misunderstandings about **diverse approval**

Diverse approval is **X** not the identification of content that is approved of by a majority...

...it is the identification of content that is approved of by diverse audiences.

Diverse approval **X** does not require (epistemic) agreement...

...only positive affect, potentially in spite of substantive disagreement.

Diverse approval is **X** not a complete solution to bridging...

...it is a particularly promising signal, (with downsides that must be mitigated).

Why?

How?

Does it work?

Why?

How?

Does it work?



The White House

@WhiteHouse



Seniors are getting the biggest increase in their Social Security checks in 10 years through President Biden's leadership.



Readers added context they thought people might want to know

Seniors will receive a large Social Security benefit increase due to the annual cost of living adjustment, which is based on the inflation rate. President Nixon in 1972 signed into law automatic benefit adjustments tied to the Consumer Price Index. Pub. L. No. 92-336 (1972). [ssa.gov/history/1972am...](https://ssa.gov/history/1972am...)

Do you find this helpful?

Rate it

Context is written by people who use Twitter, and appears when rated helpful by others. [Find out more.](#)

**MODEL**

$$\hat{r}_{un} = \mu + i_u + i_n + f_u \cdot f_n$$

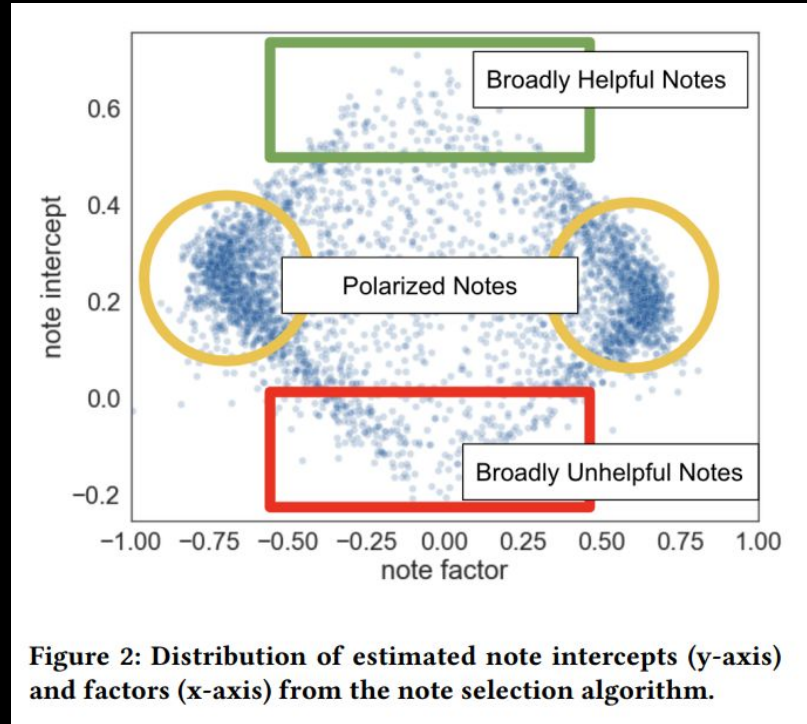


**MODEL**

$$\hat{r}_{un} = \mu + i_u + i_n + f_u \cdot f_n$$

$$\sum_{r_{un}} (r_{un} - \hat{r}_{un})^2 + \lambda_i(i_u^2 + i_n^2 + \mu^2) + \lambda_f(\|f_u\|^2 + \|f_n\|^2)$$

**LOSS FUNCTION**



Why?

How?

Does it work?

Why?

How?

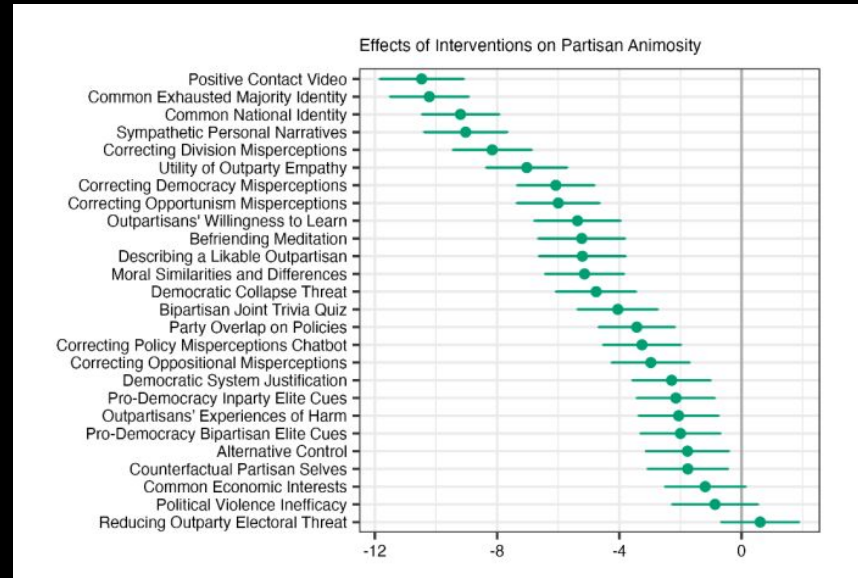
Does it work?



# Media Effects

(psychology; Strengthening Democracy Challenge)

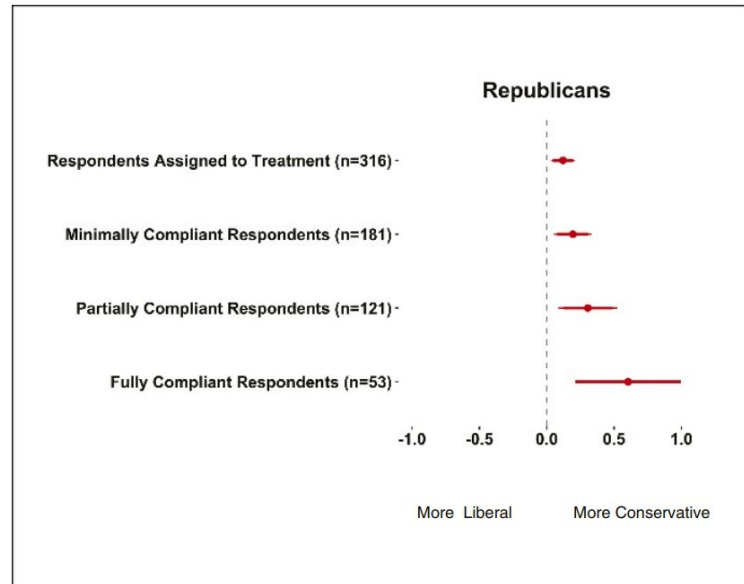
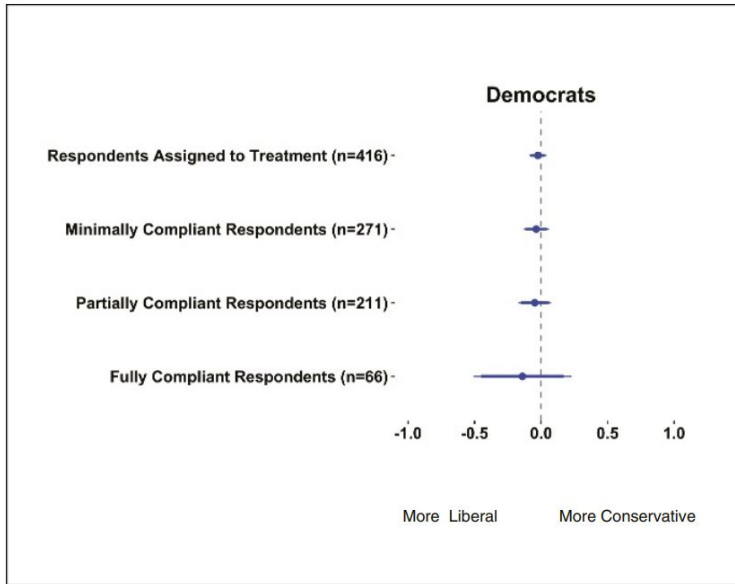
- mere-exposure effect
- illusory truth effect
- experience effects
- anchoring
- learning new facts
- discovering new interests
- pressure to conform



→ Megastudy identifying effective interventions to strengthen Americans' democratic attitudes, Voelkel et al. (2023)

# 'Mere' diversity is insufficient.

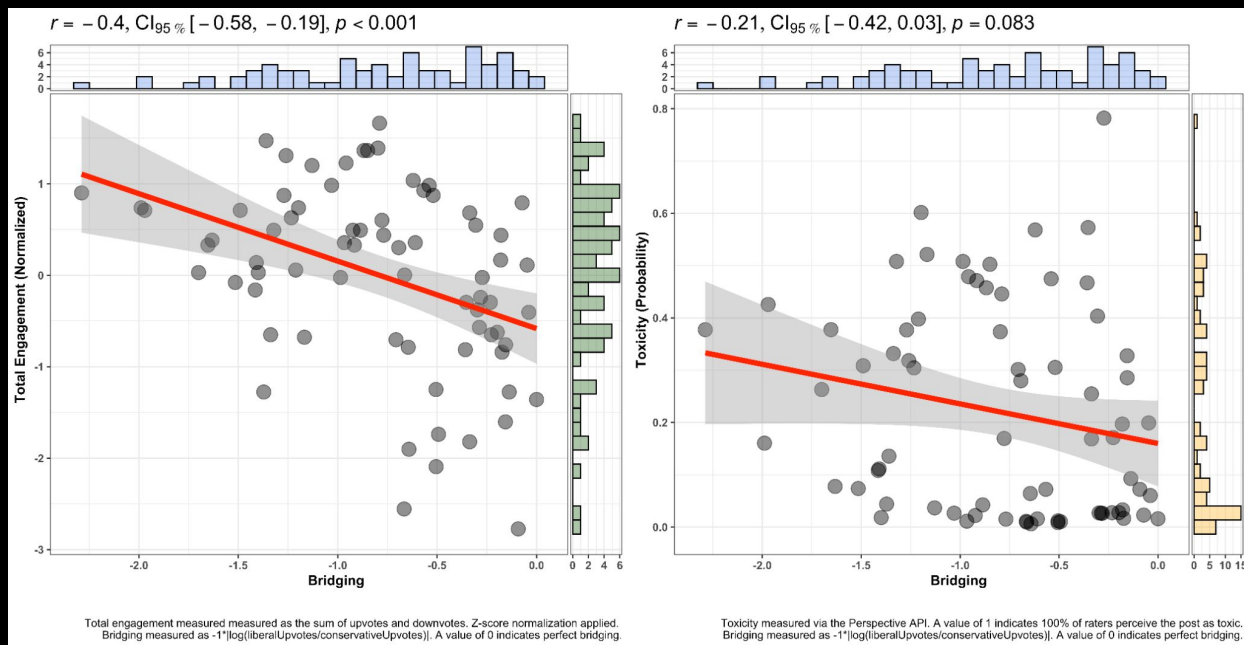
(Chris Bail)



→ Exposure to opposing views on social media can increase political polarization, Bail et al. (2018)

# Lab Evidence

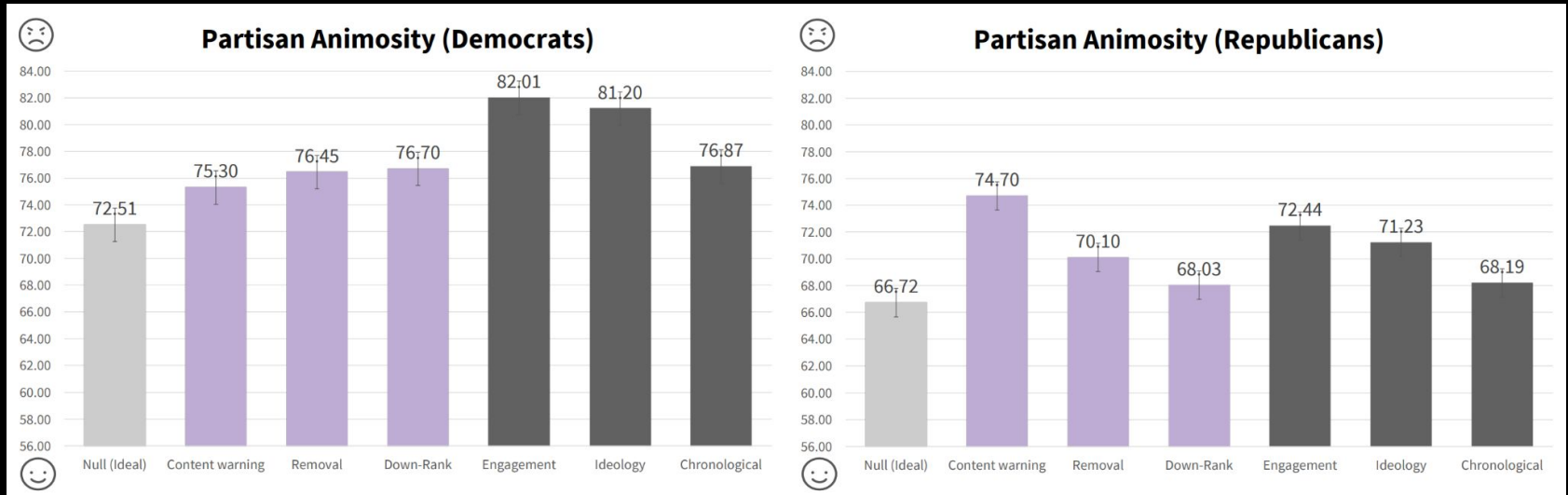
(Jason Burton, Max Planck Institute for Human Development)





# More Lab Evidence

(Stanford)



# YourView

(Tim van Gelder)

The screenshot shows the homepage of the YourView website. At the top left is the 'YourView' logo. To the right are 'SIGN IN' and 'REGISTER' buttons. Below the logo is a navigation menu with 'HOME', 'ABOUT', 'FORUMS', 'VOLUNTEER', and 'DONATE'. A search bar with 'Google Custom Search' is on the right. A banner for 'New to YourView?' with a 'Register now' link is below the menu. The main content area features a post titled 'Fairfax 2013 Federal Election' with a thumbnail of the Australian flag. The post statistics are: 41 TOPICS, 17713 VOTES, 1579 COMMENTS, and 1248 PARTICIPANTS. The post text states: 'This forum provides a unique opportunity to express your views on a broad range of major election-related issues, see how similar they are to political parties and others, and contribute to the collective wisdom, an alternative to standard opinion poll results. The Forum is a collaboration between Fairfax Media and YourView Australia, and will run throughout the election period.'

## The carbon price should be repealed

BY JAMES WRIGHT | 827 VOTES | 110 COMMENTS

### The problem

**Greenhouse gases** in the Earth's atmosphere trap heat on the surface. Greenhouse gas levels are rapidly increasing due to emissions from human activities, particularly **carbon dioxide (CO2) from fossil fuel burning**. CO2 in the atmosphere is like radioactive waste in that much of today's pollution will hang around for a very long time. Greenhouse gas pollution now far outstrips natural influences on climate, causing the Earth to accumulate heat. This problem is called **global warming**, and it is already causing serious impacts such as worsening heatwaves, droughts, and floods.

Many countries have introduced policies, including carbon prices, to reduce greenhouse gas emissions. However, present global efforts are far from adequate to achieve the **rapid emissions cuts urgently needed** to prevent dangerous levels of

The screenshot shows a forum discussion titled 'Views (110)'. At the top right are 'Columns' and 'Sort By' dropdown menus. The discussion is split into two columns: 'FOR (28)' and 'AGAINST (82)'. Each column has a 'Comment title...' and 'Comment body...' input field, and a 'Submit' button. The 'FOR' column shows a comment by Sean Baker (14.8) titled 'a start in the right direction' with 1 reply and 8 likes. The 'AGAINST' column shows a comment by bkenry (80.6) titled 'Lets take the arguments for repeal one by one' with 13 likes. The 'AGAINST' column also shows a comment by -pat (13.5) titled 'Its a vote grabbing wealth redistribution' with 13 likes. The 'AGAINST' column also shows a comment by -lindypenguin (47.1) titled 'You have to start somewhere' with 13 likes.

## Views (110)

Display  Columns  Sort By

### FOR (28)

Comment title... 0/50  
Comment body...  
Submit

**a start in the right direction** -Sean Baker (14.8)  
Look, a lot of the resistance comes from a willful blindness to the science. Once that is acknowledged and we can move on, the question becomes how do we collectively address reducing our carbon emissions, and encouraging ourselves to do more with less. This is a good start.  
2 years ago 1 reply 8

### AGAINST (82)

Comment title... 0/50  
Comment body...  
Submit

**Lets take the arguments for repeal one by one** -bkenry (80.6)  
Tax hurts businesses and households  
Yes, initially it probably will. Especially 'trade exposed' industries. However, the damage from things like electricity price increases is marginal compared to the increases we have all endured over the last few years which have had nothing to do with the tax. Except for those businesses which pay the tax directly, the carbon tax is actually one of the smallest cost  
... more  
3 years ago 13

**Its a vote grabbing wealth redistribution** -pat (13.5)  
Whilst I support the underlying ethos of moving away from fossil fuels to clean energy and that government has a major role to play in creating a need for businesses and consumers to change.  
I can not support it in its current form. I think the best way forward is to start again. Incorporate a no disadvantage to Australian exporters, an impost on imports where a carbon tax or equivalent has not been levied

**You have to start somewhere** -lindypenguin (47.1)  
To avoid dangerous climate change Australia will have to do 4 things: reduce domestic emissions, lobby for international action, become a

→ Cultivating Deliberation for Democracy, Tim van Gelder (2020)  
→ Public Wisdom, Tim van Gelder (2020)

# Polis

(Computational Democracy Project)

## How divisive was the conversation?



Consensus statements

Divisive statements

% Agreed % Disagreed % Passed % Didn't vote

STATEMENT	OVERALL 177	A 152	B 25
5 I think we will have to accept a lower level of transparency than we are used to, but that this will be worth it given the gains AI offers.	19% 64% 15% (57)	15% 66% 17% (45)	33% 58% 8% (12)
53 We need more civil society-led initiatives - independent from corporate influence - to lead inclusive, democratic conversations about AI.	89% 0% 10% (58)	86% 0% 13% (45)	100% 0% 0% (13)
66 It will be helpful to study GAI systems in terms of how their training data include or exclude different coalitions of people.	80% 9% 10% (65)	79% 10% 10% (49)	81% 6% 12% (16)

# X / Twitter

(Community Notes)



The White House  @WhiteHouse

Seniors are getting the biggest increase in their Social Security checks in 10 years through President Biden's leadership.

 Readers added context they thought people might want to know

Seniors will receive a large Social Security benefit increase due to the annual cost of living adjustment, which is based on the inflation rate. President Nixon in 1972 signed into law automatic benefit adjustments tied to the Consumer Price Index. Pub. L. No. 92-336 (1972). [ssa.gov/history/1972am...](https://ssa.gov/history/1972am...)

Do you find this helpful?

Context is written by people who use Twitter, and appears when rated helpful by others. [Find out more.](#)

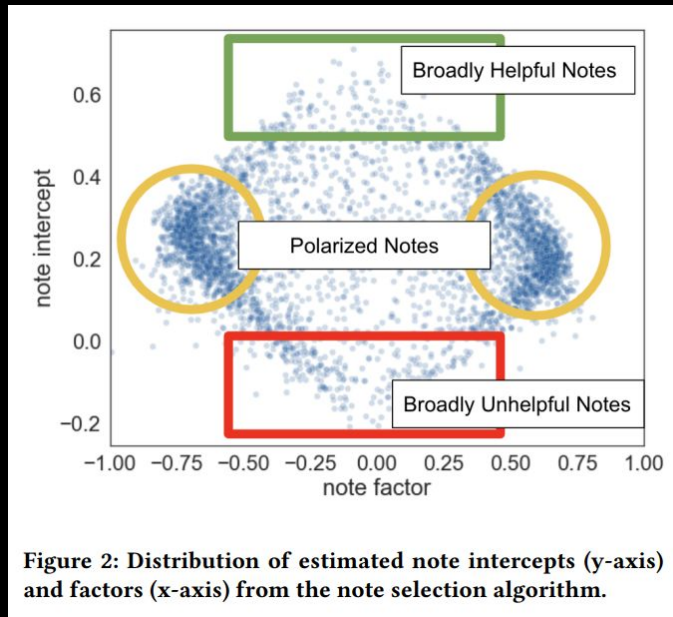
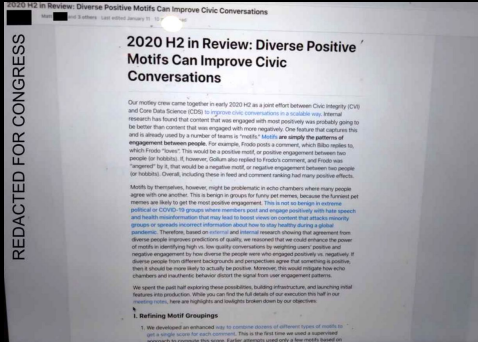


Figure 2: Distribution of estimated note intercepts (y-axis) and factors (x-axis) from the note selection algorithm.

# Facebook

(Facebook Papers)



Comments with high positive motifs from a highly diverse audience — relative to those with high positive motifs from a *non*-diverse audience — were:

- **10.9x less likely to be reported**
- **3.4-5.2x lower rates of violations than comments**

At the user-level, people with a reputation for getting positive motifs from more diverse audiences were:

- **much less likely to violate community standards,**
- **much less likely to be reported by other users for being fake, bullying, hate, or violence incitement**

“Thus, reputations for garnering positive motifs are especially informative of the likely value of content when enhanced by the diversity of people engaging with that content.”



→ bridging.systems  
→ lukethorburn.com

Thanks to the many collaborators & mentors who have contributed to or informed this work →



This work is conducted in partnership with the GETTING-Plurality Research Network housed within the Allen Lab for Democracy Renovation at Harvard University.

This work was supported by UK Research and Innovation [grant number EP/S023356/1], in the UKRI Centre for Doctoral Training in Safe and Trusted Artificial Intelligence (safeandtrustedai.org), King's College London.



Maria Polukarov  
KCL



Carmine Ventre  
KCL



Aviv Ovadya  
Berkman Klein



Jonathan Stray  
CHAI



Smitha Milli  
Cornell Tech



Helena Puig Larrauri  
BuildUp



Seth Lazar  
ANU



Emillie de Keulenaar  
UN

...