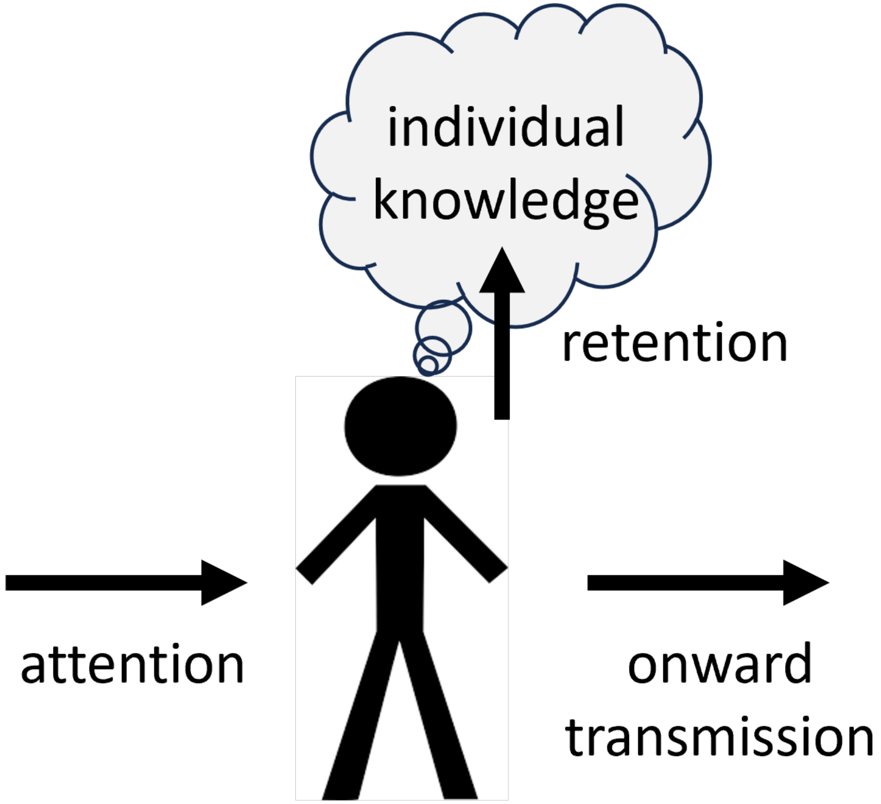
“Social Attention and Information Transmission”

David McAdams

Stanford GSB Economics Seminar

October 15, 2025

A Model of Socially-Acquired Knowledge



Advancing an Economic-Epidemiology Agenda

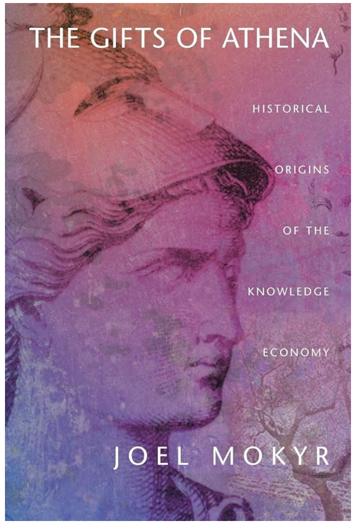
CREATION of information	**DIFFUSION** of information	**IMPACT** of information

Advancing an Economic-Epidemiology Agenda

CREATION of information	**DIFFUSION** of information	**IMPACT** of information
Mokyr (2002)		Mokyr (2002)

*"The crucial [factor behind Europe's Industrial-Era surge in technological and economic growth was] a small group of at most a few thousand people who formed **a creative community based on the exchange of knowledge**."*

-- from "The Gifts of Athena" (2002)



Advancing an Economic-Epidemiology Agenda

CREATION of information	**DIFFUSION** of information	**IMPACT** of information
Mokyr (2002)	Han, **Hirshleifer**, Walden (2023) [J. Finance: "Visibility Bias in Tranmission … "]	Mokyr (2002)

"A missing chapter in our understanding of finance consists of the social processes that shape economic thinking and behavior … [including] social transmission bias, ***the systematic directional modification of ideas or signals as they pass from person to person****."*

-- from 2020 AFA Presidential Address

Advancing an Economic-Epidemiology Agenda

CREATION of information	**DIFFUSION** of information	**IMPACT** of information

McAdams et al (2019)
[PLoS Biology: "Resistance Diagnostics ... "]

"We can potentially turn the Darwinian tables on antibiotic-resistant bacteria ... We shouldn't settle for a post-antibiotic fate."

-- from Duke press release "Could Better Tests Help Reverse the Rise of Superbugs?"

Sam Brown [microbial ecology]
GA Tech Biological Sciences

Marc Lipsitch [epidemiology]
Harvard Chan

Advancing an Economic-Epidemiology Agenda

CREATION of information	**DIFFUSION** of information	**IMPACT** of information

Kranton & McAdams (2024)
[AEJ Micro: "… Information Markets"]

McAdams et al (2019)
[PLoS Biology: "Resistance Diagnostics … "]

McAdams & Song (2025)
[TE: "Adoption Epidemics …"]

Jackson, Malladi & McAdams (2022,2025)
[PNAS: "Learning through the grapevine …"
WP: "Rational disagreement …"]

Kranton & McAdams (2024)
[AEJ Micro: "… Information Markets"]

McAdams & Song (2025)
[TE: "Adoption Epidemics …"]

Advancing an Economic-Epidemiology Agenda

CREATION of information	**DIFFUSION** of information	**IMPACT** of information

Kranton & McAdams (2024)
[AEJ Micro: "… Information Markets"]

McAdams et al (2019)
[PLoS Biology: "Resistance Diagnostics … "]

THIS PAPER

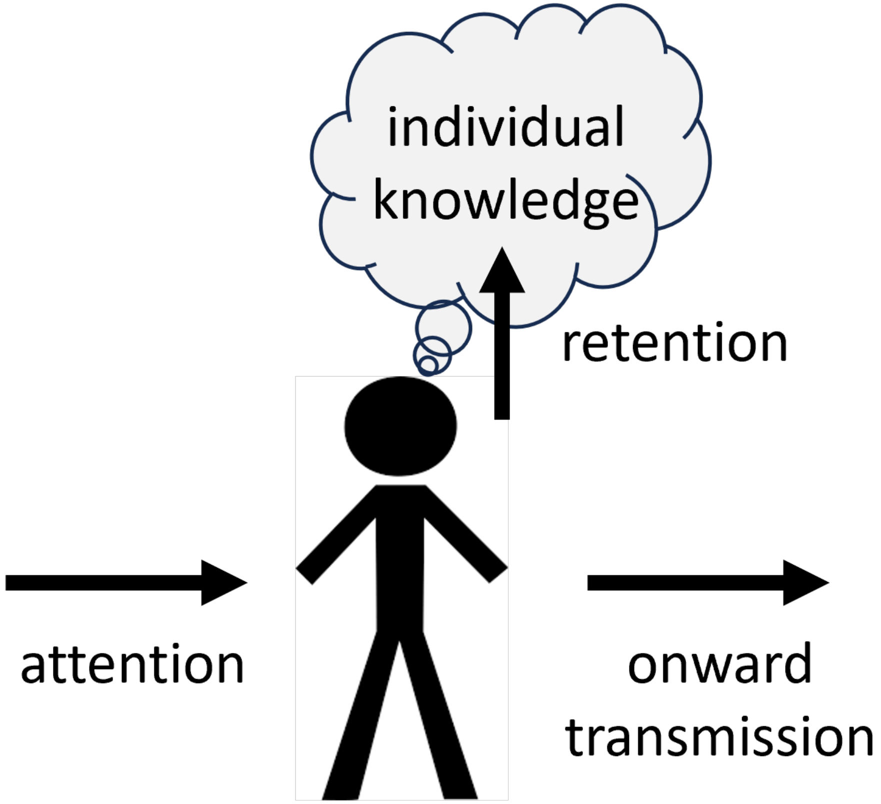
McAdams & Song (2025)
[TE: "Adoption Epidemics …"]

Jackson, Malladi & McAdams (2022,2025)
[PNAS: "Learning through the grapevine …"
WP: "Rational disagreement …"]

Kranton & McAdams (2024)
[AEJ Micro: "… Information Markets"]

McAdams & Song (2025)
[TE: "Adoption Epidemics …"]

A Model of Socially-Acquired Knowledge



Plan for the Talk

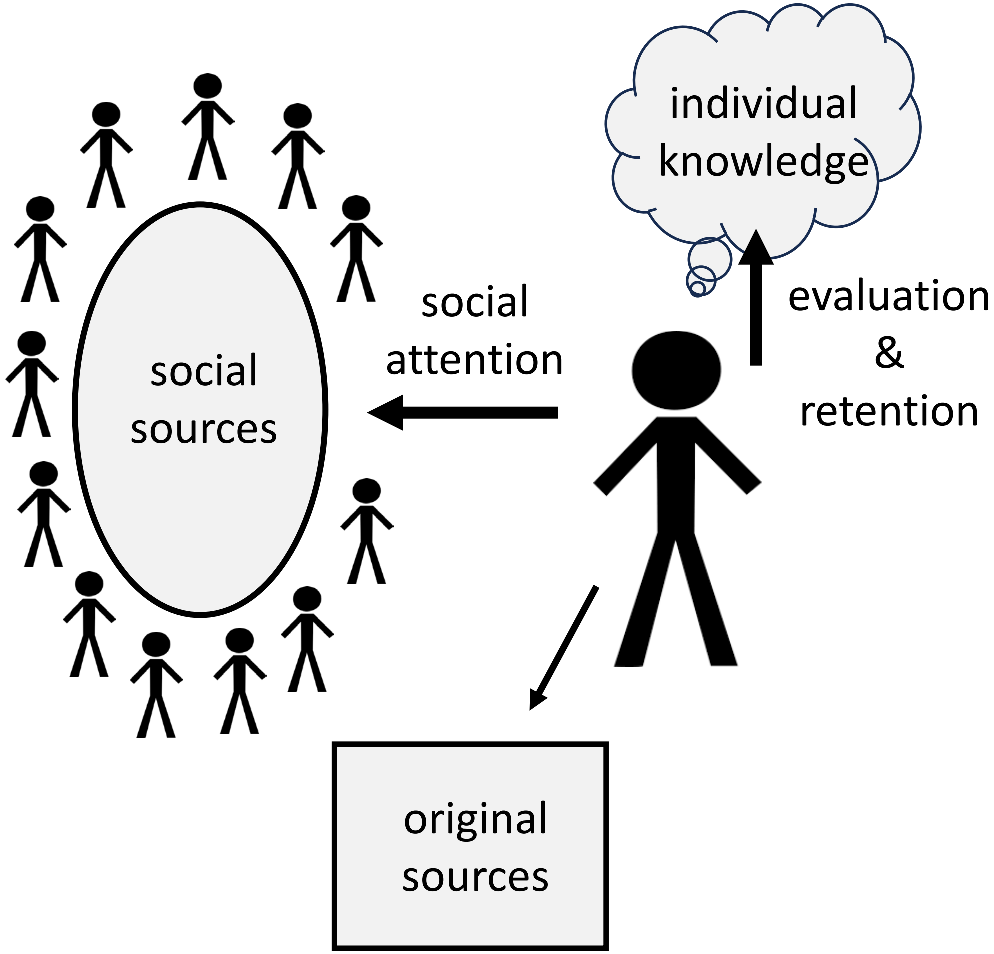
Benchmark Model

Baseline Analysis

Attractive-Story Extension

Partisan-Agent Extension

A Model of Socially-Acquired Knowledge



- **attention:** each agent *i* allocates attention across information source(s)
- **evaluation:** when exposed to a piece of information ("story"), agent *i* gets a private signal about the story …
- **retention:** … and decides whether to retain the story as knowledge
- **onward transmission:** other agents who pay attention to *i* may be exposed to any story in *i*'s store of knowledge

Model: Stories & Agents

Stories

– exogenous pool of “original stories”

– each original story is true or false

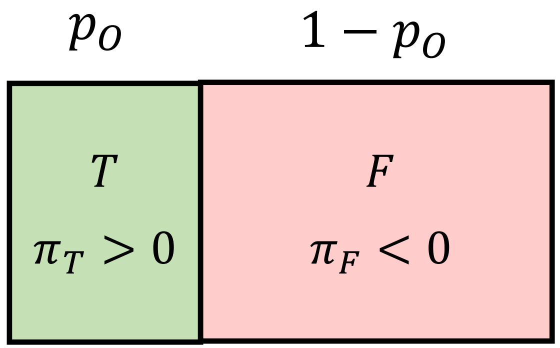
- $\omega = T$ if true; $\omega = F$ if false
- fraction p_O of original stories are true

Agents

– unit-mass population of agents

– an agent who retains a story gets knowing payoff

- $\pi_T > 0$ if true
- $\pi_F < 0$ if false



original stories
(**Pool O**)

p_O

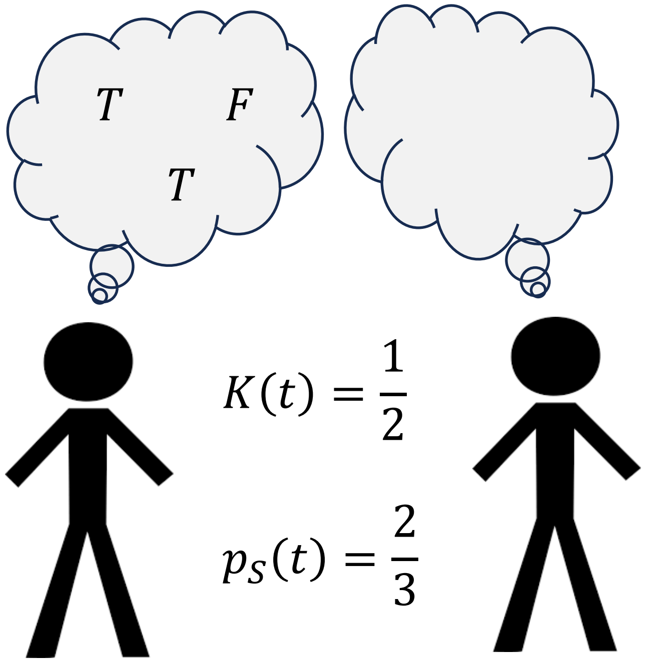
Model: Individual and Collective Knowledge

- $\boldsymbol{K}^i(t)$: individual knowledge
- $\biguplus_i \boldsymbol{K}^i(t)$: collective knowledge
  - contains multiple copies of stories known by multiple agents
  - referred to as "**Pool S**," mnemonic for "socially-shared pool"

- $N_m(t)$: mass of agents who know $m \geq 0$ stories
- $K(t) = 1 - N_0(t)$: "knowledge prevalence"

- $M_T(t)$, $M_F(t)$: mass of true and false stories in Pool S

- $p_S(t) \equiv \frac{M_T(t)}{M_T(t)+M_F(t)}$: "social veracity"



Model: Attention to Original and Social Sources

– $\lambda_O^i(t)$: rate that agent i encounters a randomly-selected original story

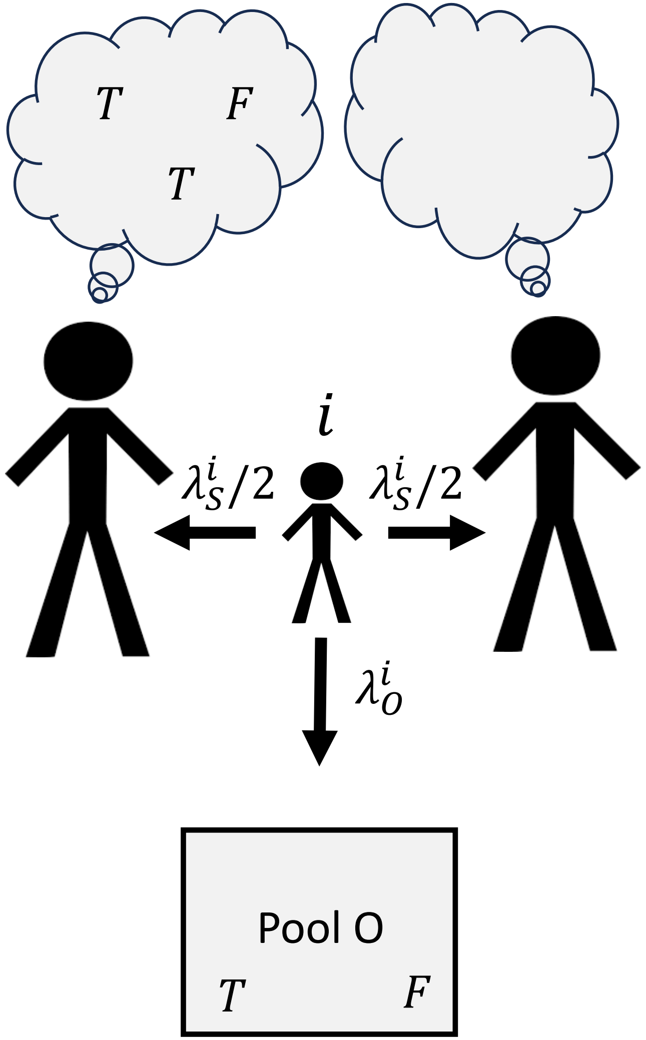
  - fraction p_O of original stories are true

– $\lambda_S^i(t)$: rate that agent i has social interactions

  - during a social interaction with agent j, i is exposed to a randomly-selected story from j's knowledge
  - $\lambda_S^i(t)K(t)$: rate that agent i has a social interaction with someone who has a story to share
  - fraction $p_S(t)$ of socially-encountered stories are true

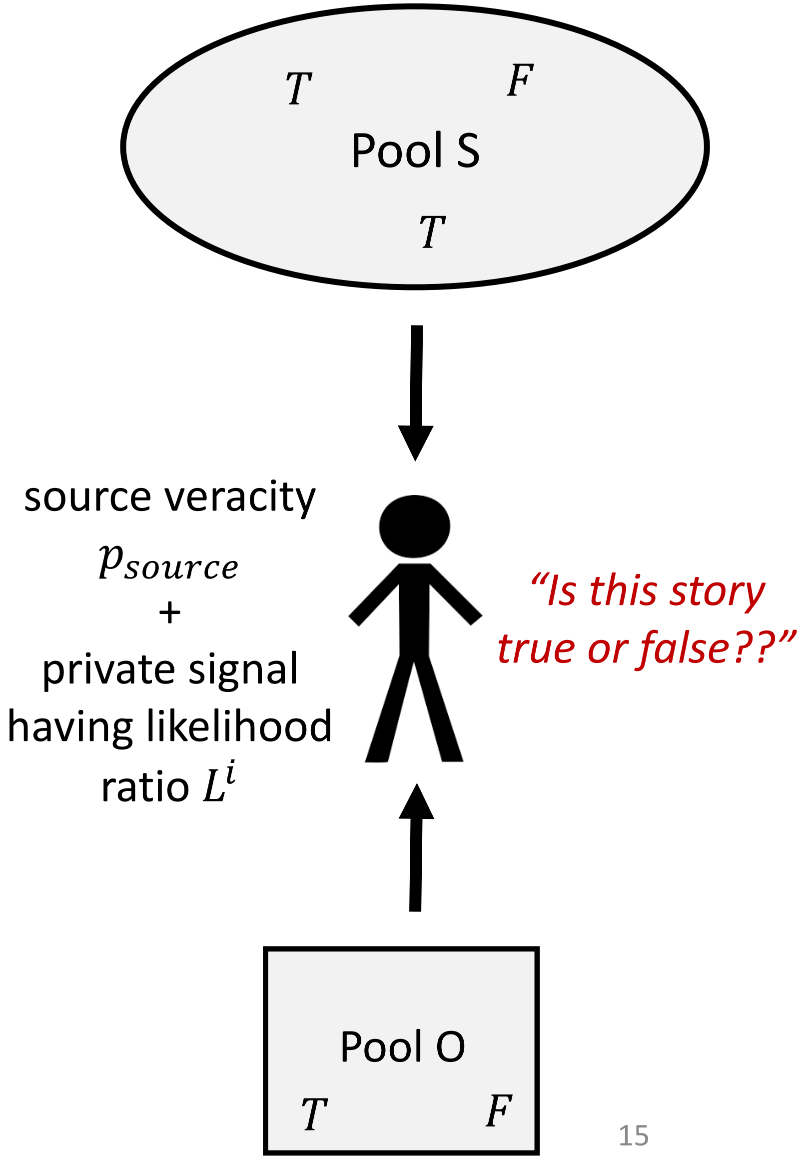
– each agent chooses $(\lambda_O^i(t), \lambda_S^i(t))$ subject to

  - $\underbrace{\lambda_O^i(t) + \lambda_S^i(t) \le \bar{\lambda};}_{\text{“attention budget”}} \quad \underbrace{\lambda_O^i(t) \ge \underline{\lambda}_O > 0; \quad \lambda_S^i(t) \ge 0}_{\text{“automatic/minimal exposures”}}$



Model: Private Signals and Story Evaluation

- upon encountering a story, each agent observes
  - the source of the story (Pool O or Pool S)
  - private signal $s^i \in [0, \infty)$ with continuous pdf $g(s^i|\omega)$
- … and Bayes-updates their belief

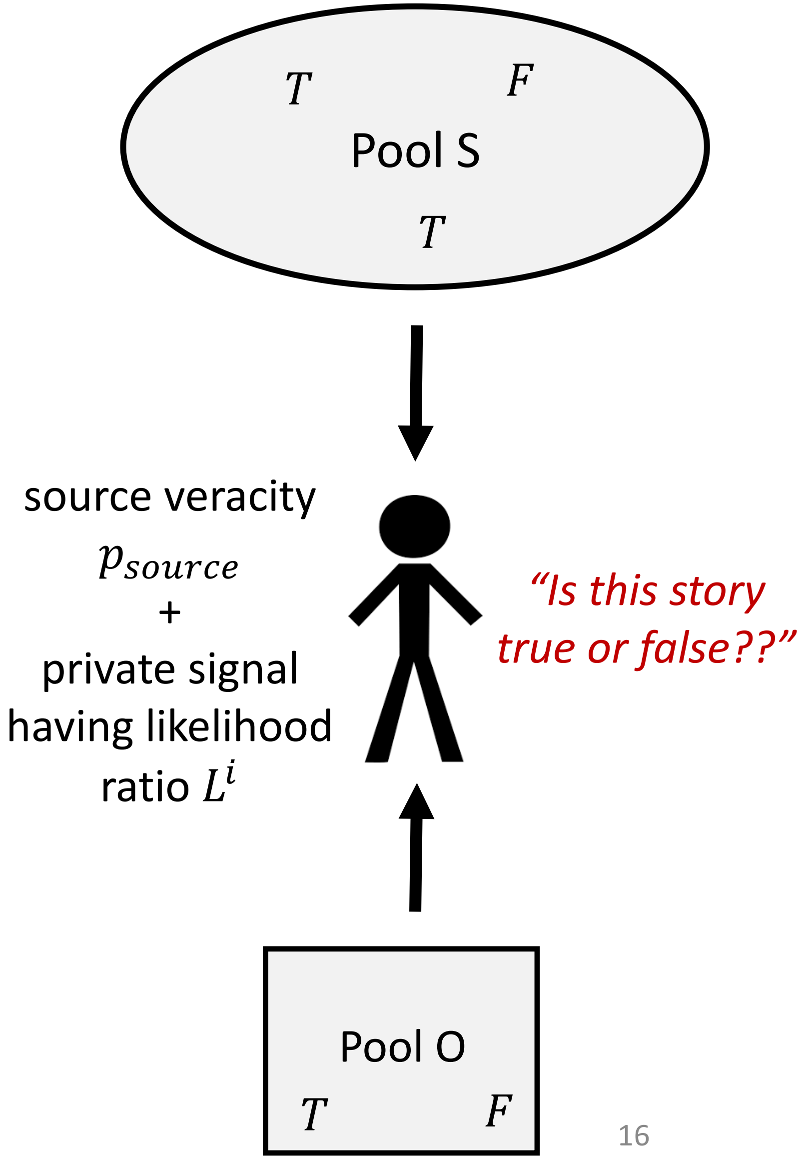


Model: Private Signals and Story Evaluation

- upon encountering a story, each agent observes
  - the source of the story (Pool O or Pool S)
  - private signal $s^i \in [0, \infty)$ with continuous pdf $g(s^i|\omega)$
- ... and Bayes-updates their belief to p^i, given by

$$\frac{p^i}{1-p^i} = \frac{p_{source}}{1-p_{source}} \times (\text{signal likelihood ratio } L^i)$$

- assume: likelihood ratio $L^i(s^i) = \frac{g(s^i|\omega=T)}{g(s^i|\omega=F)}$ strictly increasing, with $L^i(0) = 0$ and $\lim_{s^i \to \infty} L^i(s^i) = \infty$
- henceforth, I will view $L^i \in [0, \infty)$ as the signal



Model: Story Retention

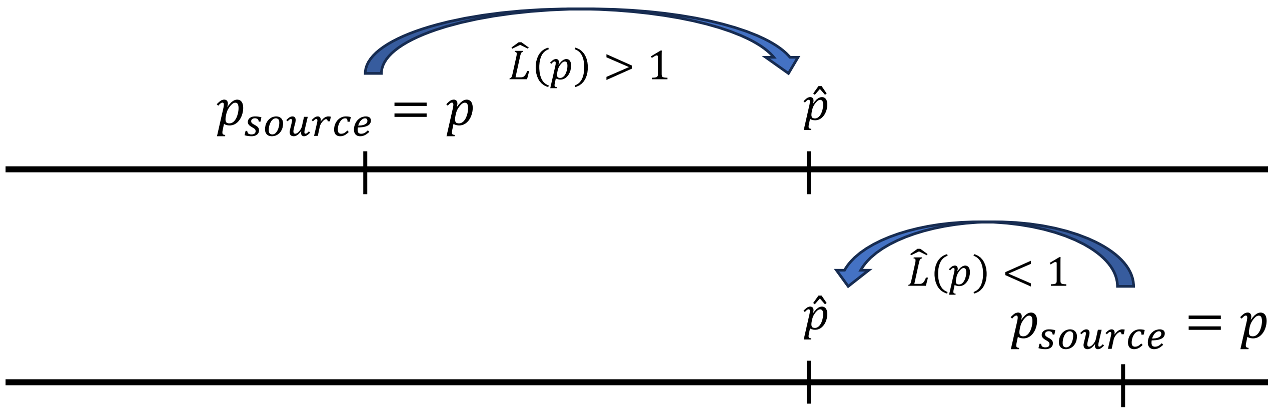
– agent i's expected payoff from retaining a story is

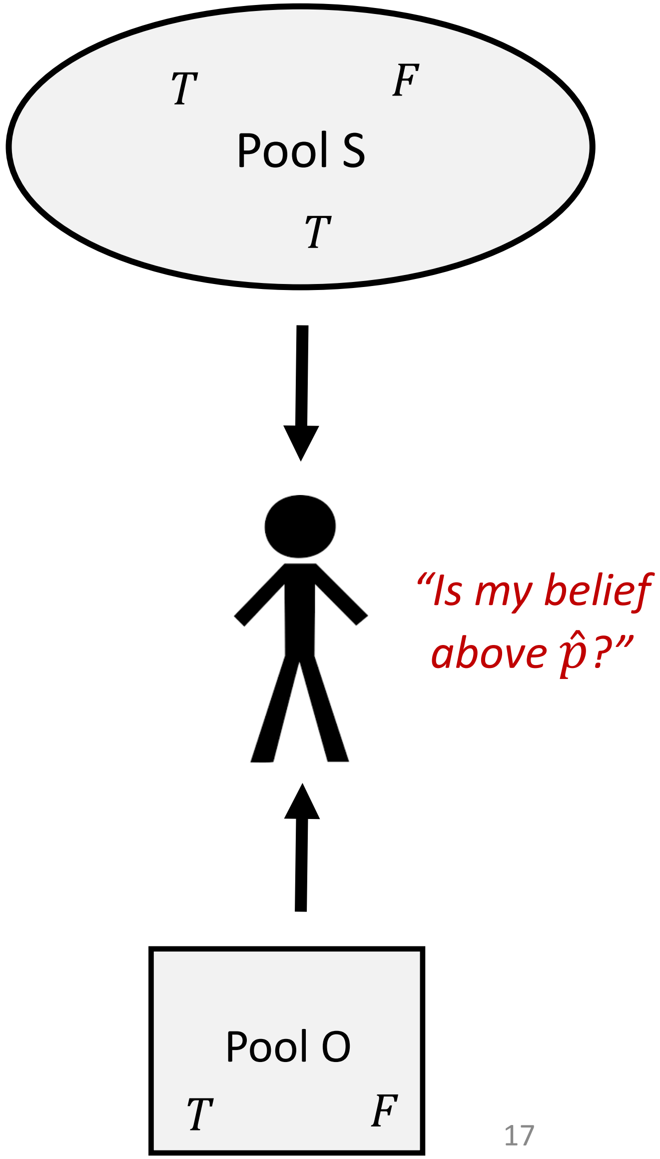
$$p^i \pi_T + (1 + p^i)\pi_F$$

– retention is optimal whenever $p^i \geq \hat{p} \equiv \frac{-\pi_F}{\pi_T - \pi_F}$

- $\hat{p}$ is the threshold belief for story retention

– let $\hat{L}(p)$ be the threshold signal for retention to be optimal, from a source with veracity $p_{source} = p$





Model: Forgetting, Initial Condition, and Eqm

Forgetting

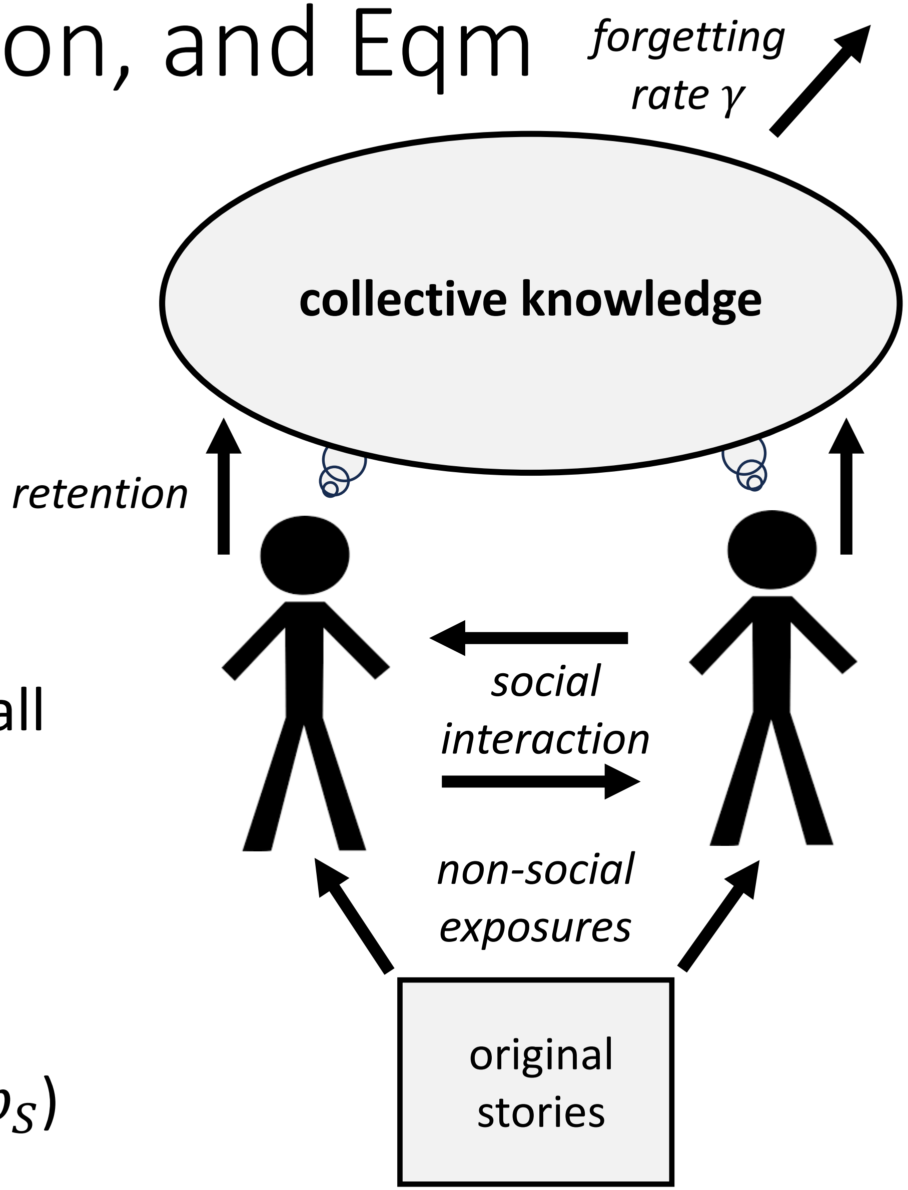
- each agent forgets each story they know independently at rate γ
- **assumption for the talk:** $\gamma = 0$

Initial condition

- **assumption:** at time 0, agents know nothing at all

Equilibrium

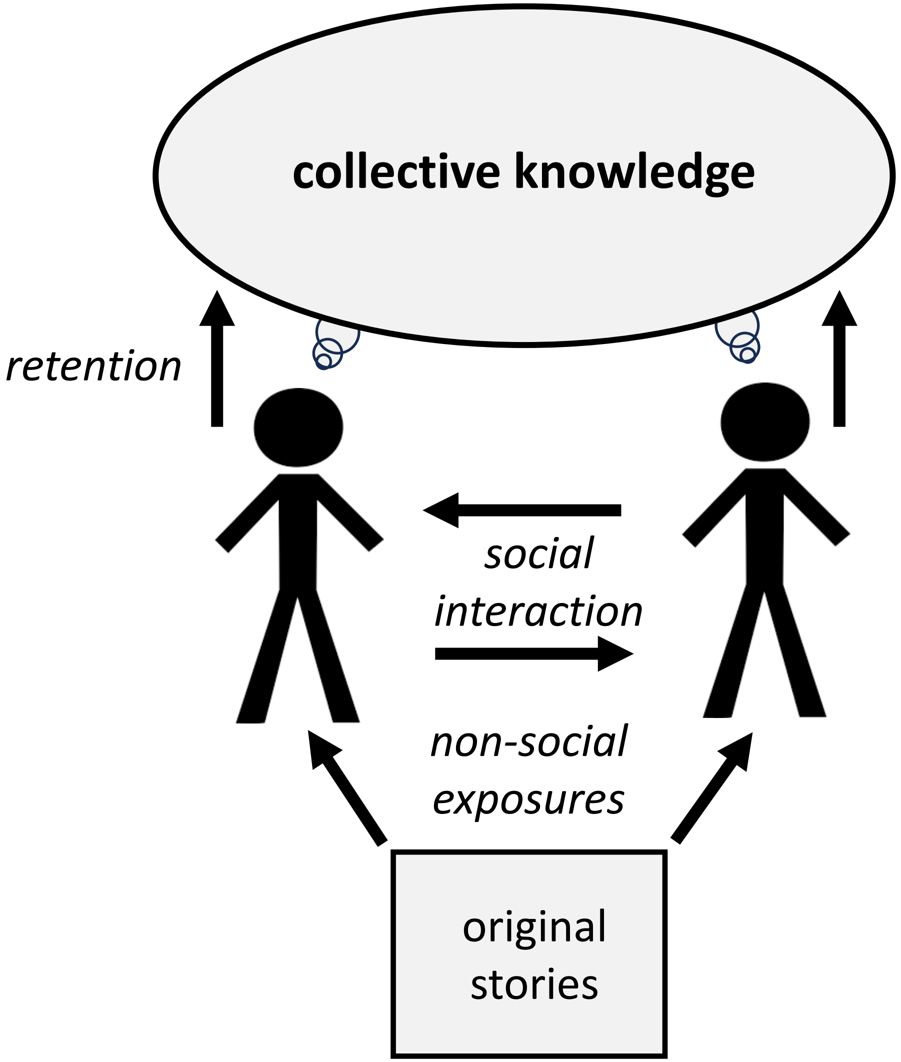
- strategies = attention & retention
- key objects: individual knowledge (N_0, N_1, …); collective knowledge (M_T, M_F); social veracity (p_S)



Model Summary

BENCHMARK MODEL

- stories are true or false
- homogeneous agents
- agents allocate attention across original & social sources
- agents decide what stories to retain as knowledge



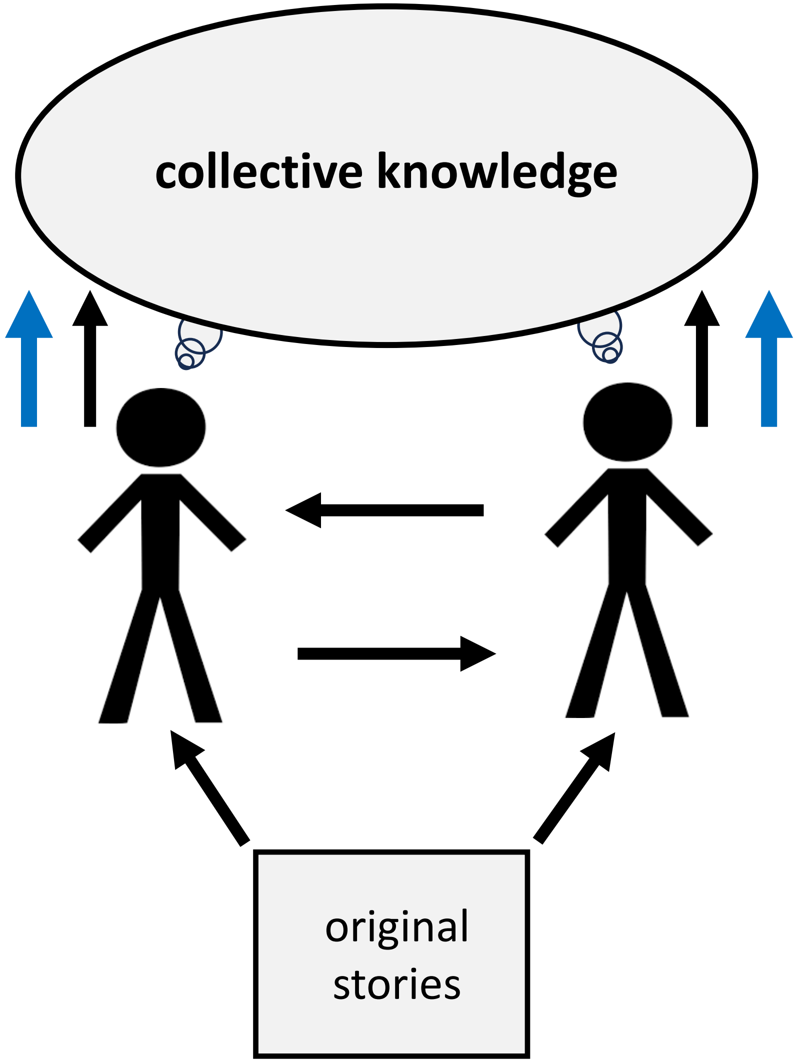
Model Summary

BENCHMARK MODEL

- stories are true or false
- homogeneous agents
- agents allocate attention across original & social sources
- agents decide what stories to retain as knowledge

ATTRACTIVE-STORIES EXTENSION

- stories are also attractive or neutral
- agents enjoy knowing attractive stories, even if false



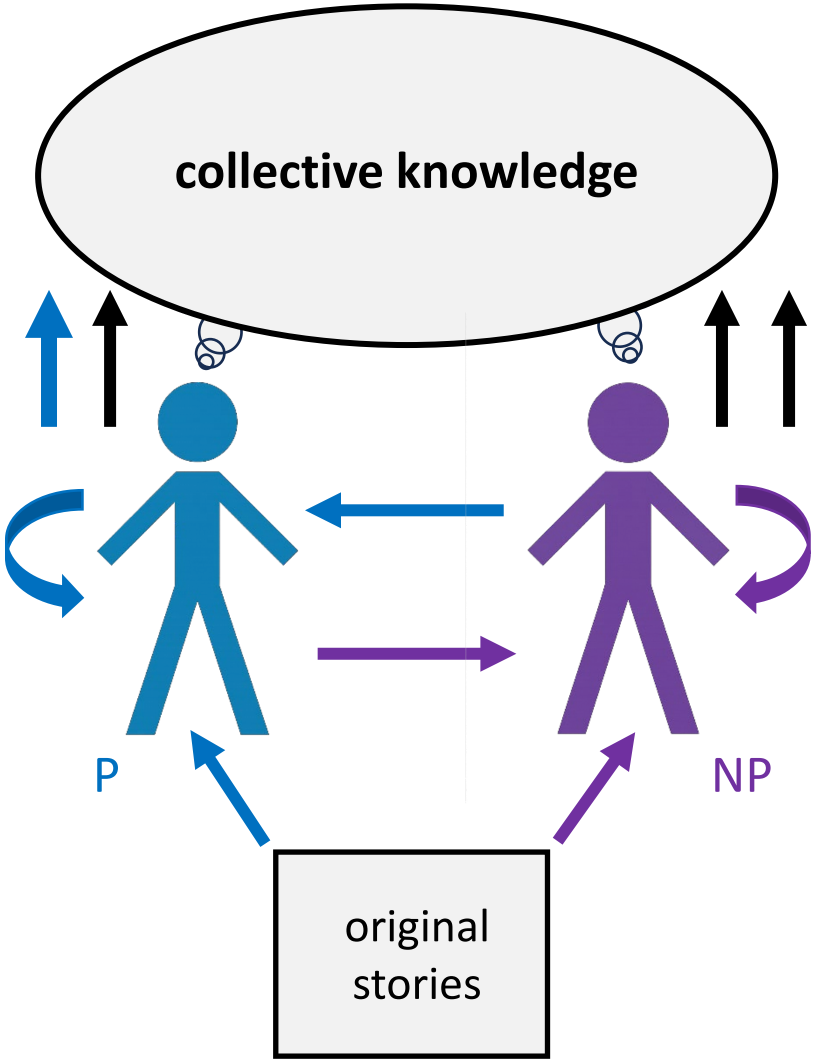
Model Summary

BENCHMARK MODEL

- stories are true or false
- homogeneous agents
- agents allocate attention across original & social sources
- agents decide what stories to retain as knowledge

ATTRACTIVE-STORIES EXTENSION

- stories are also attractive or neutral
- agents enjoy knowing attractive stories, even if false



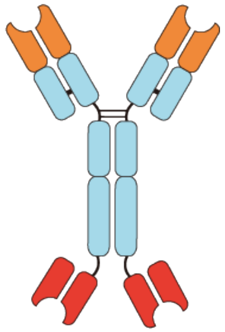
PARTISAN-AGENT EXTENSION

- agents are “partisan” (P) or “non-partisan” (NP)
- only partisans care about attractiveness, now interpreted as “aligned with partisan view”
- agents allocate attention across original sources, own type, and other type
- endogenous equilibrium mixing

Examples: Socially-Transmitted Info

Grandma's recipe

The Big Lie



Binding reagents

2018 study: More widely-shared recipes tend to be both tastier and healthier.

2024 poll: 38% of US adults believed that Biden's 2020 victory was not legitimate in Jan 2024, up from 31% in Dec 2021.

2015 call to action by world's leading chemical-probe researchers: Biomedical researchers persistently fail to use objectively superior binding reagents.

Why does social transmission seem to "work" in promoting high-quality recipes, but not truth in politics or research methods in biomedicine?

Key Findings: Benchmark Model

Strategic complementarity

- as people "listen" more to others, stories are filtered more intensively via social transmission, which increases the payoff to social attention

Features of the equilibrium trajectory

- **cold start**: initially, agents only attend to original sources
- **monotone**: social veracity is weakly increasing over time
- **sociality promotes quality content:** enabling people to have more social interactions strictly increases equilibrium social veracity

Key Findings: Attractive-Story Extension

Endogenous negative correlation

– even if “attractive” stories are inherently equally likely to be true, the attractive stories *in social circulation* are less likely to be true

Crowding out

– widespread sharing of attractive stories “crowds out” other stories, resulting in neutral stories also having lower veracity than in the benchmark

Key Findings: Partisan-Agent Extension

Non-partisan agents (NPs) only listen to one another

– intuitively: *“NPs like me are more trustworthy sources”*

Partisan agents (Ps) form a similar bubble if *either* inherent story veracity is high *or* they have “extreme” preferences for partisan content

– intuitively: *“why should I listen to those guys??”*

But if inherent veracity is low and partisanship is “moderate,” Ps split their attention in equilibrium

Plan for the Talk

Benchmark Model

Baseline Analysis

- attention dynamics
- knowledge dynamics
- veracity dynamics

Attractive-Story Extension

Partisan-Agent Extension

Attention Allocation

Let $E\Pi(p)$ be the expected benefit due to encountering a story from a source with veracity p

Marginal value of attention to Pool O: $E\Pi(p_O) > 0$

Marginal value of attention to Pool S: $K(t)E\Pi(p_S(t))$

Since agents initially know nothing, $K(t) \approx 0$ at all times $t \approx 0$

→ The equilibrium trajectory begins in an **Isolated Phase** in which agents only pay attention to original sources, namely, $\lambda_O^i = \bar{\lambda}$ for all i

Story Retention

Let $\psi(t)$ be the rate at which each agent optimally retains stories

During the Isolated Phase,

$$\psi(t) = \psi_O \equiv \bar{\lambda}(p_O \sigma_T(p_O) + (1 - p_O)\sigma_F(p_O))$$

where

$$\sigma_T(p_O) \equiv \Pr\left(L^i > \hat{L}(p_O) \mid \omega = T\right)$$

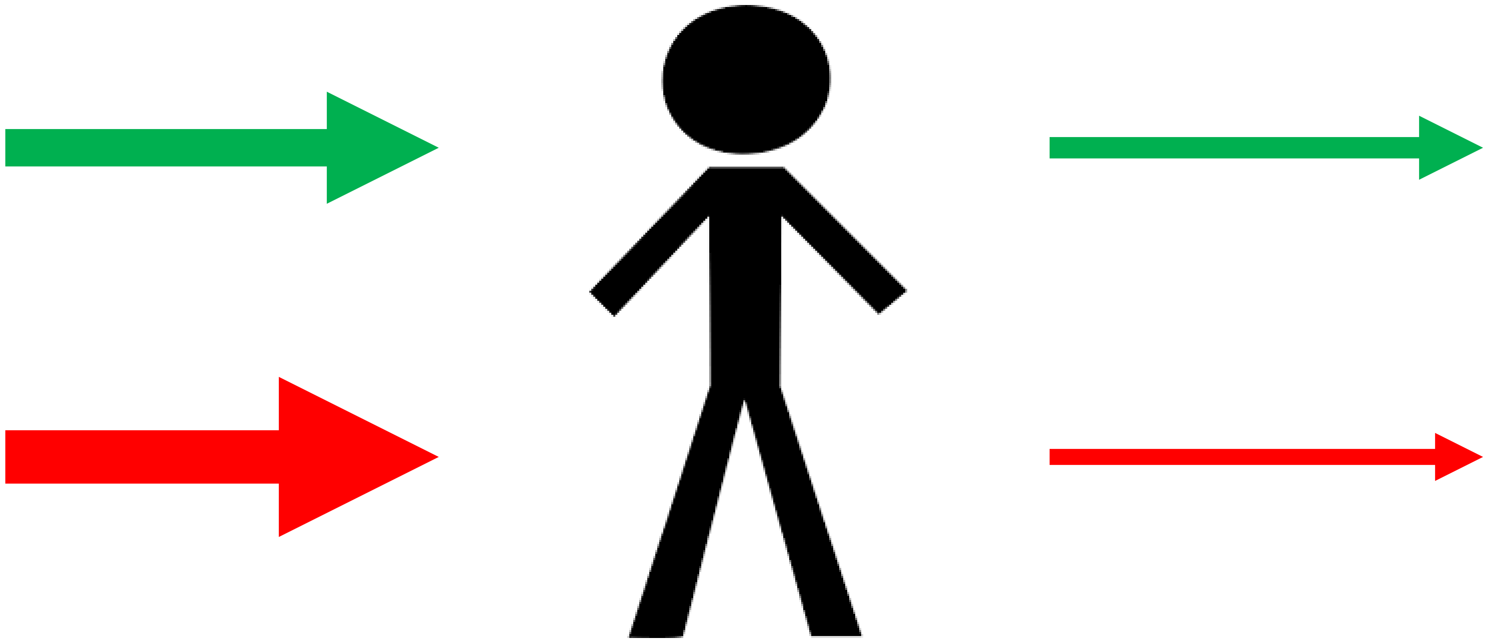
$$\sigma_F(p_O) \equiv \Pr\left(L^i > \hat{L}(p_O) \mid \omega = F\right)$$

are the *retention likelihoods* for true or false stories from Pool O.

As shorthand, define $\sigma_{OT} \equiv \sigma_T(p_O)$ and $\sigma_{OF} \equiv \sigma_F(p_O)$

Inflow of Newly-Retained Stories

Example with $p_O < \frac{1}{2}$ *and* $\hat{p} = \frac{1}{2}$



Retained stories are mostly true

Retained volume is low

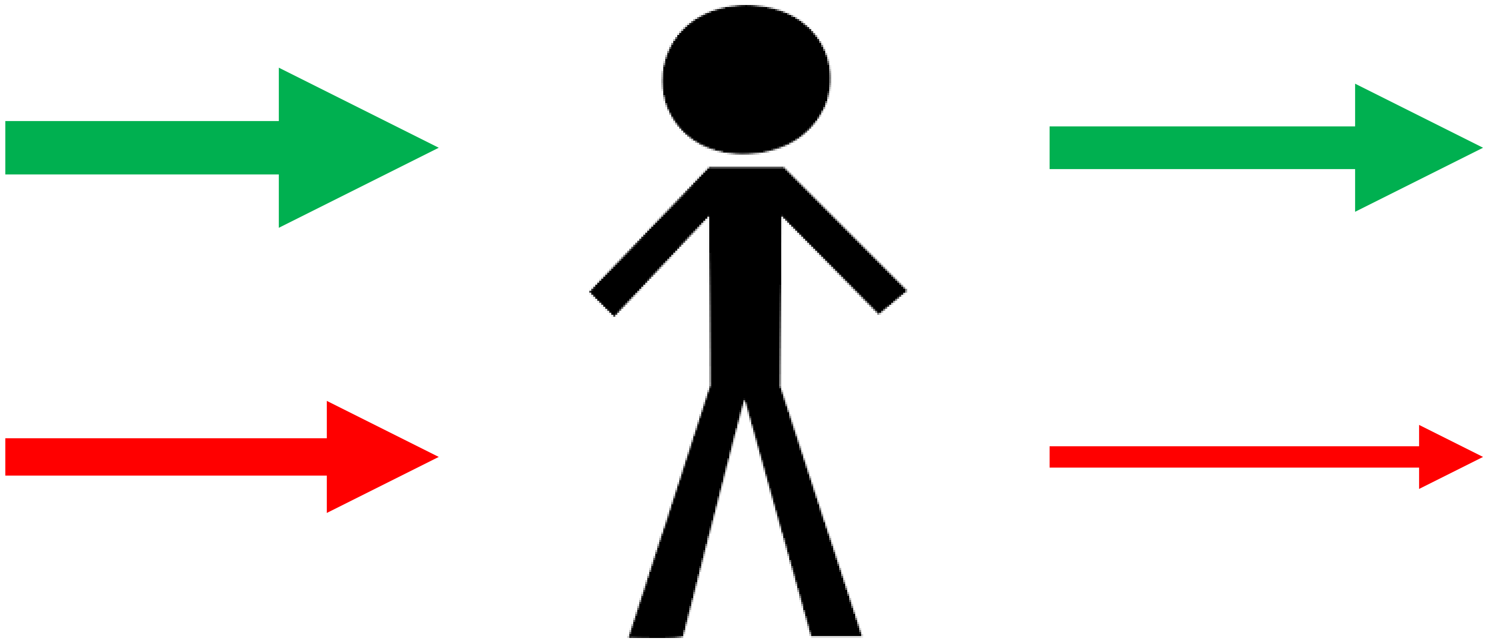
from original sources

into knowledge

$p_{in}(p_O; \hat{p}) > \frac{1}{2}$

Inflow of Newly-Retained Stories

Example with $p_O > \frac{1}{2}$ *and* $\hat{p} = \frac{1}{2}$



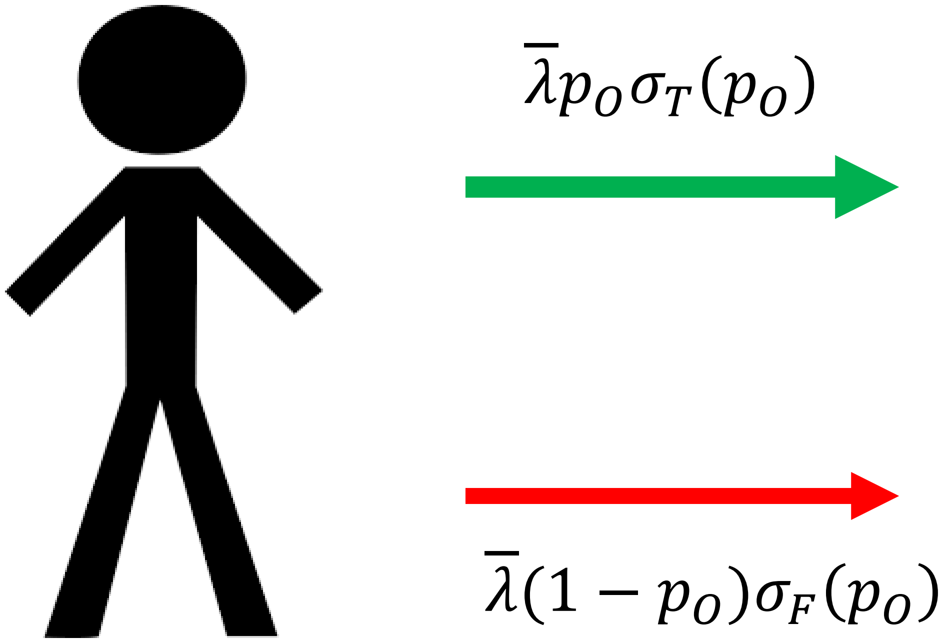
Retained stories even more likely true

Retained volume is high

from original sources

into knowledge $p_{in}(p_O; \hat{p}) > p_O$

"Inflow Veracity"



Stories entering the knowledge pool from original sources have **"inflow veracity"** $p_{in}(p_O)$, where

$$p_{in}(p_O) \equiv \frac{p_O \sigma_T(p_O)}{p_O \sigma_T(p_O) + (1 - p_O)\sigma_F(p_O)}$$

“Inflow Veracity”

Inflow Veracity Lemma

i. $p_{in}(p) > \max\{p, \hat{p}\}$

ii. $p_{in}(p)$ strictly increasing and continuous in p

iii. $p_{in}(p)$ strictly increasing and continuous in $\hat{p}$

Intuition: Retained stories are positively selected, only retained if the agent’s signal is good enough for Bayes-updated belief to exceed $\hat{p}$

Attentional Incentives during Isolated Phase

Pool O

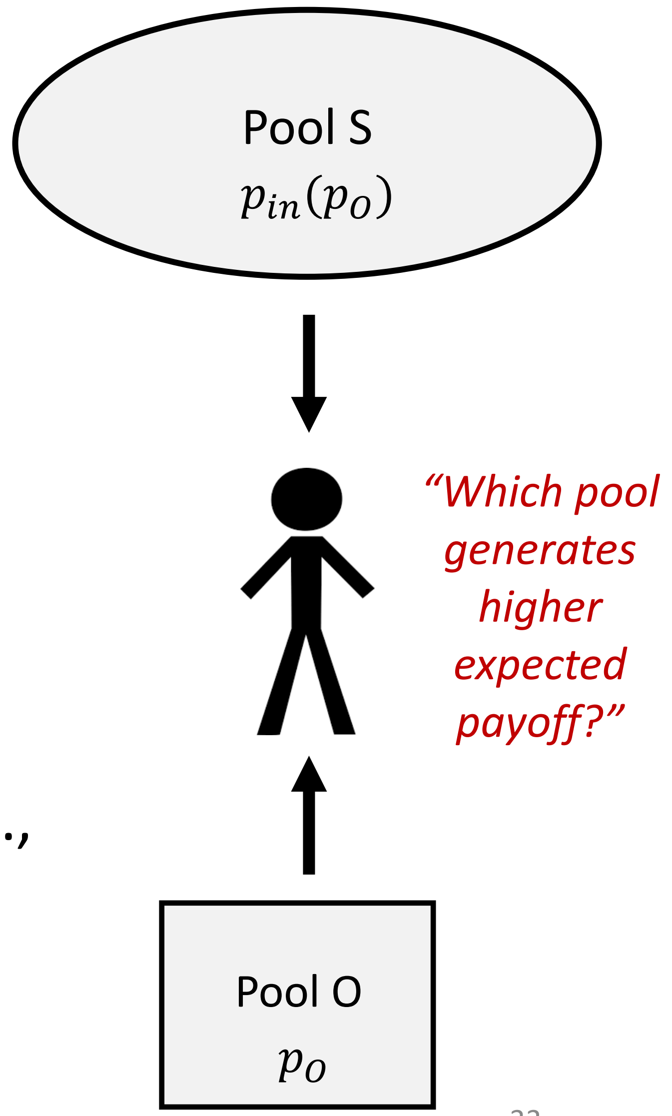
- *advantage*: always see a story
- *disadvantage*: lower veracity p_O

Pool S

- *advantage*: higher veracity $p_{in}(p_O)$
- *disadvantage*: fraction $1 - K(t)$ of social interactions are duds, with no story exposure

Isolation continues so long as $E\Pi(p_O) > K(t)E\Pi(p_{in}(p_O))$, i.e.,

$$\boxed{K(t) < \hat{K} \equiv \frac{E\Pi(p_O)}{E\Pi(p_{in}(p_O))}}$$



Plan for the Talk

Benchmark Model

Baseline Analysis

- attention dynamics
- **knowledge dynamics**
- veracity dynamics

Attractive-Story Extension

Partisan-Agent Extension

Individual-Knowledge Dynamics

Intuition: # stories known at time t is Poisson distributed …

Individual-Knowledge Lemma

i. $N_m(t) = \frac{\mu(t)^m}{m!}\exp(-\mu(t))$ *for all* $m = 0,1,2,\ldots,$ *where*

$$\mu(t) \equiv \int_0^t \psi(s)\exp\big(-\gamma(t-s)\big)\mathrm{d}s$$

ii. *If* $\lim_{t\to\infty}\psi(t) = \psi^\infty$, *then* $\lim_{t\to\infty}K(t) = K^\infty$, *where*

$$\boxed{K^\infty = 1 - \exp\left(\frac{-\psi^\infty}{\gamma}\right)}$$

Individual-Knowledge Dynamics

Individual-Knowledge Lemma

i. $N_m(t) = \frac{\mu(t)^m}{m!}\exp(-\mu(t))$ *for all* $m = 0,1,2,\ldots,$ *where*

$$\mu(t) \equiv \int_0^t \psi(s)\exp\big(-\gamma(t-s)\big)\mathrm{d}s$$

ii. *If* $\lim_{t\to\infty}\psi(t) = \psi^\infty$, *then* $\lim_{t\to\infty} K(t) = K^\infty$, *where*

$$\boxed{K^\infty = 1 - \exp\left(\frac{-\psi^\infty}{\gamma}\right)}$$

Intuition: *# stories known at time t is Poisson distributed …*

Given that $\gamma = 0$ *[assumption for talk],* $K(t)\uparrow$ *and* $K^\infty = 1.$

Transition: Isolated Phase → Social Phase

Let $\hat{t}$ be the time at which $K(\hat{t}) = \widehat{K}$.

$$p_S^*(t) = p_{in}(p_O)$$
$$\lambda_S^*(t) = 0$$
$$\lambda_0^*(t) = \overline{\lambda}$$
$$\psi^*(t) = \psi_O$$
$$K(t) < \widehat{K}$$

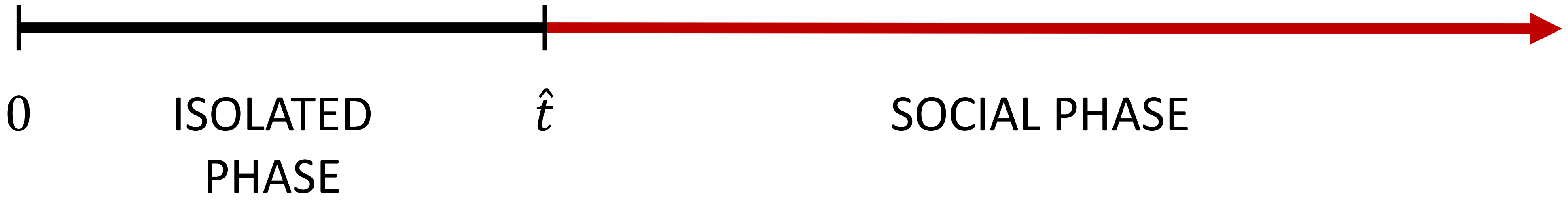
$$p_S^*(\hat{t}) = p_{in}(p_O)$$
$$\lambda_S^*(\hat{t}\,+) > 0$$
$$\lambda_0^*(\hat{t}\,+) = \overline{\lambda} - \lambda_S^*(\hat{t}\,+)$$

$\psi^*(\hat{t})$ has ambiguous discontinuity at $\hat{t}$

$K(\hat{t}) = \widehat{K}$ has ambiguous kink at $\hat{t}$

attention shifts to social sources

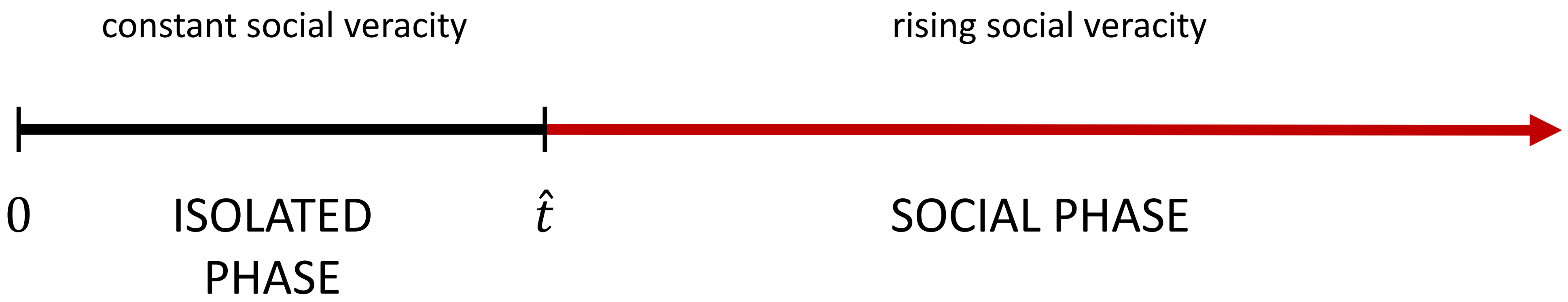
rate of knowledge acquisition may rise or fall, depending on details



Social Phase: Virtuous Circle

Social transmission drives a virtuous circle

- social filtering → rising social veracity [details next]
- rising social veracity + rising knowledge prevalence → rising benefit of social interaction



Plan for the Talk

Benchmark Model

Baseline Analysis

- attention dynamics
- knowledge dynamics
- **veracity dynamics**

Attractive-Story Extension

Partisan-Agent Extension

Story-Pool Dynamics

Recall that social veracity $p_S(t) = \frac{M_T(t)}{\Sigma M(t)}$, where $\Sigma M(t) = M_T(t) + M_F(t)$

The story-pool state $(M_T(t), M_F(t))$ is governed by diff-equations

$$M_T'(t) = \underbrace{\lambda_O(t) p_O \sigma_{OT}}_{\textit{inflow from Pool O}} + \underbrace{\lambda_S(t) K(t) p_S(t) \sigma_T\big(p_S(t)\big)}_{\textit{inflow from Pool S}} - \underbrace{\gamma M_T(t)}_{\textit{outflow = zero}}$$

$$M_F'(t) = \lambda_O(t)(1 - p_O)\sigma_{OF} + \lambda_S(t) K(t)(1 - p_S(t))\sigma_F\big(p_S(t)\big) - \gamma M_F(t)$$

Social-Veracity Dynamics

Using that $M_T(t) = p_S(t)\, \Sigma M(t)$, we can rearrange these diff-eqs as

$$\frac{M_T'(t)}{M_T(t)} = \frac{1}{\Sigma M(t)} \left(\frac{\lambda_O(t) p_O \sigma_{OT}}{p_S(t)} \right) + \lambda_S(t) K(t) \sigma_T\big(p_S(t)\big)$$

$$\frac{M_F'(t)}{M_F(t)} = \frac{1}{\Sigma M(t)} \left(\frac{\lambda_O(t)(1 - p_O)\sigma_{OF}}{1 - p_S(t)} \right) + \lambda_S(t) K(t) \sigma_F\big(p_S(t)\big)$$

Finally, note that $p_S'(t) > 0$ if and only if $\frac{M_T'(t)}{M_T(t)} > \frac{M_F'(t)}{M_F(t)}$

Social-Veracity Dynamics

$$\frac{M_T'(t)}{M_T(t)} - \frac{M_F'(t)}{M_F(t)} = IN\big(t; p_S(t)\big) + SF\big(t; p_S(t)\big)$$, where

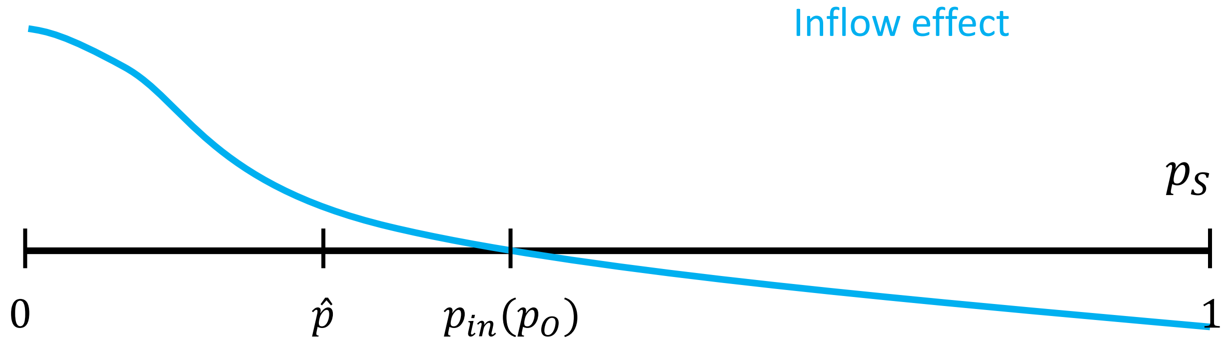
Inflow effect: $$IN(t; p_S) \equiv \lambda_O(t)\left(\sigma_{OT}\frac{p_O}{p_S} - \sigma_{OF}\frac{1-p_O}{1-p_S}\right)$$

Social-filtering effect: $$SF(t; p_S) \equiv \overbrace{\lambda_S}^{\textit{social attention}}\underbrace{(t)K}_{\textit{knowledge prevalence}}\overbrace{(t)(\sigma_T(p_S) - \sigma_F(p_S))}^{\textit{extra likelihood true stories retained}}$$

Social-Veracity Dynamics

Inflow and Social-Filtering Effects Lemma. At all times t:

i. $IN(t; p_S)$ is decreasing in p_S and zero when $p_S = p_{in}(p_O)$



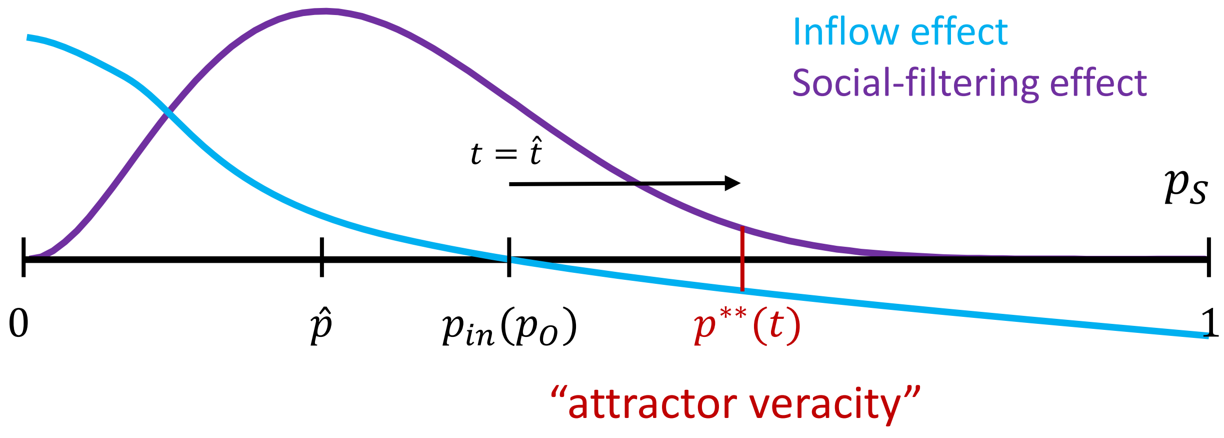
Social-Veracity Dynamics

Inflow and Social-Filtering Effects Lemma. At all times t:

i. $IN(t; p_S)$ is decreasing in p_S and zero when $p_S = p_{in}(p_O)$

ii. $SF(t; p_S) > 0$ for all $p_S \in (0,1)$ and zero at $p_S = 0,1$

iii. $SF(t; p_S)$ is quasiconcave in p_S and maximized at $p_S = \hat{p}$



"Attractor Veracity"

Define the "attractor veracity" $p^{**}(t)$ implicitly by the condition

$$IN(t; p^{**}(t)) + SF(t; p^{**}(t)) = 0$$

$p(t)$ always evolves toward $p^{**}(t)$

Inflow effect: $$IN(t; p_S) \equiv \lambda_O(t)\left(\sigma_{OT}\frac{p_O}{p_S} - \sigma_{OF}\frac{1-p_O}{1-p_S}\right)$$

Social-filtering effect: $$SF(t; p_S) \equiv \overbrace{\lambda_S}^{\textit{social attention}}\underbrace{(t)K}_{\textit{knowledge prevalence}}\overbrace{(t)(\sigma_T(p_S) - \sigma_F(p_S))}^{\textit{extra likelihood true stories retained}}$$

Social-Veracity Dynamics

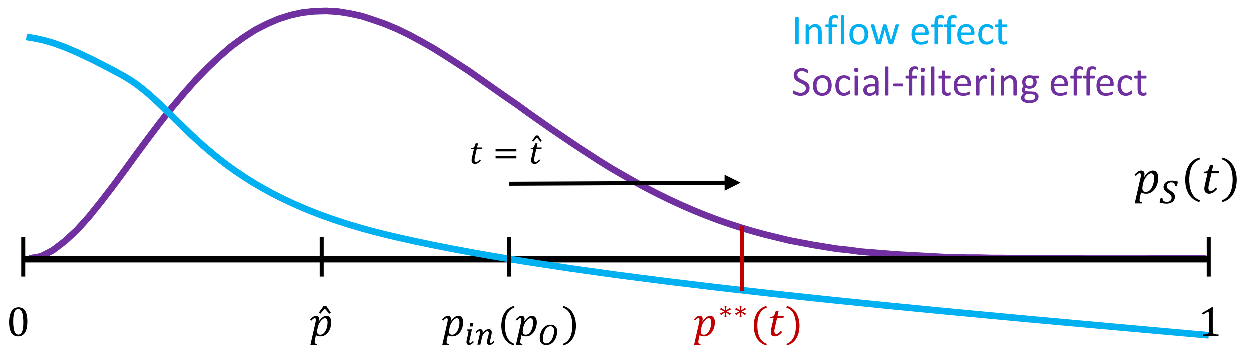
Inflow and Social-Filtering Effects Lemma. At all times t:

i. $IN(t; p_S)$ is decreasing in p_S and zero when $p_S = p_{in}(p_O)$

ii. $SF(t; p_S) > 0$ for all $p_S \in (0,1)$ and zero at $p_S = 0,1$

iii. $SF(t; p_S)$ is quasiconcave in p_S and maximized at $p_S = \hat{p}$

iv. $SF(t; p_S)$ and hence $p^{**}(t)$ is increasing in $K(t)$ and $\lambda_S(t)$



Monotonicity of the Attractor Veracity

Let $p^{**}(\lambda_O, \lambda_S, K)$ denote the attractor veracity, viewed as a function of agent attention to original and social sources, and knowledge prevalence

– *up until time* $\hat{t}$: $\lambda_S = 0$ and $p^{**}(t) = p_{in}(p_O)$

– *at time* $\hat{t}$: social attention jumps up to $\lambda_S = \overline{\lambda} - \underline{\lambda}_O$, causing $p^{**}(t)$ to jump up to $p^{**}\left(\underline{\lambda}_O, \overline{\lambda} - \underline{\lambda}_O, \widehat{K}\right) > p_{in}(p_O)$

– *after time* $\hat{t}$: $p^{**}(t) = p^{**}(\underline{\lambda}_O, \overline{\lambda} - \underline{\lambda}_O, K(t))$ which, given the no-forgetting assumption, is monotonically increasing to 100%

$$p^{**\infty} = p^{**}(\underline{\lambda}_O, \overline{\lambda} - \underline{\lambda}_O, 1)$$

Plan for the Talk

Benchmark Model

Baseline Analysis

Attractive-Story Extension

- modeling “attractiveness”
- impact on attention dynamics
- impact on veracity dynamics
- equilibrium welfare implications

Partisan-Agent Extension

Model: Stories & Agents

Stories

- exogenous pool of "original stories"
- each original story is (i) *true* or *false* ($\omega = T, F$) and (ii) *attractive* or *neutral* ($\chi = +, 0$)
  - fraction p_O of all original stories are true
  - fraction q_O of all original stories are attractive
- unlike truth, attractiveness is observable
- **assume:** $\Pr(T|\chi = +) = \Pr(T|\chi = 0) = p_O$, i.e., attractive stories are *inherently* equally likely to be true

	p_O	$1 - p_O$
q_O	$T+$	$F+$
$1 - q_O$	$T0$	$F0$

original stories
(Pool O)

Model: Stories & Agents

Stories

- each story is (i) unobservably *true* or *false* ($\omega = T, F$) and (ii) observably *attractive* or *neutral* ($\chi = +, 0$)

Agents

- unit-mass population of agents
- an agent who retains a story gets knowing payoff

$$\pi_T \text{ (if true)} + \pi_F \text{ (if false)} + b \text{ (if attractive)}$$

- **assume:** $b + \pi_F > 0$, i.e., agents benefit from retaining attractive stories, even if false

	p_O	$1 - p_O$
q_O	$\pi_T + b$	$\pi_F + b > 0$
$1 - q_O$	$\pi_T > 0$	$\pi_F < 0$

Pool O

Impacts of Attractiveness: During Isolated Phase

More stories retained → K_0^∞ rises

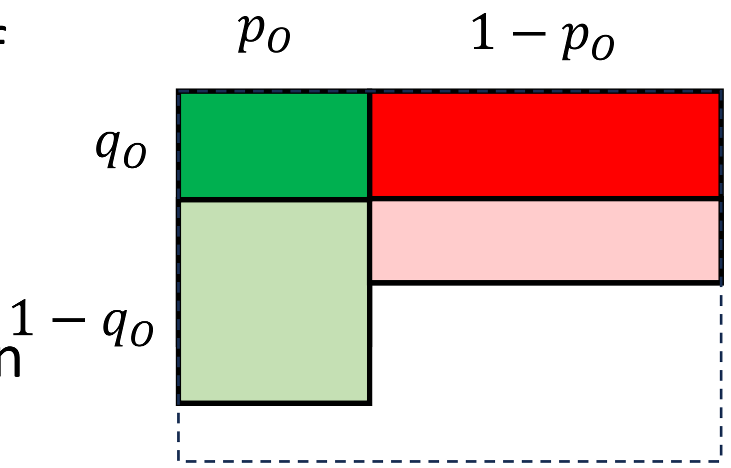
- holding all else fixed, this effect boosts the benefit of social interaction

Retained stories are disproportionately attractive

- this effect also boosts the benefit of social interaction

Retained stories have lower veracity than otherwise

- this effect reduces the benefit of social interaction



retained inflow from Pool O

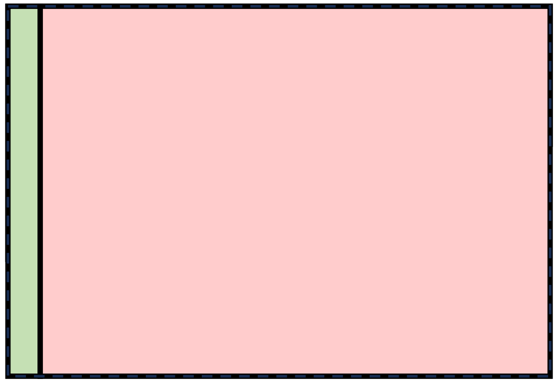
Example #1: Attractive Stories Trigger Social Filtering

Suppose that $p_0 \approx 0$, $\hat{p} = \frac{1}{2}$, and neutral stories only

Agents will retain $\approx 0\%$ of the stories that they encounter from original sources

So long as the forgetting rate $\gamma > 0$, knowledge prevalence remains $\approx 0\%$ forever

→ the Social Phase never gets off the ground!



Pool O

retained inflow from Pool O

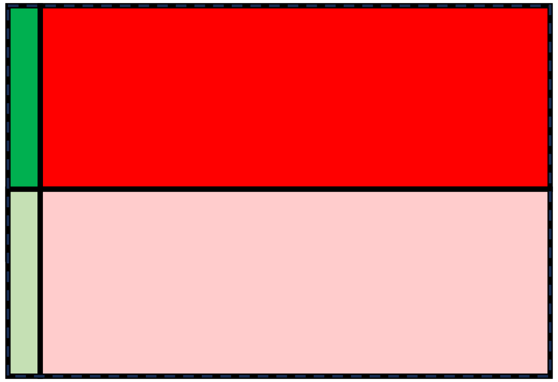
Example #1: Attractive Stories Trigger Social Filtering

Now suppose half of the stories are attractive

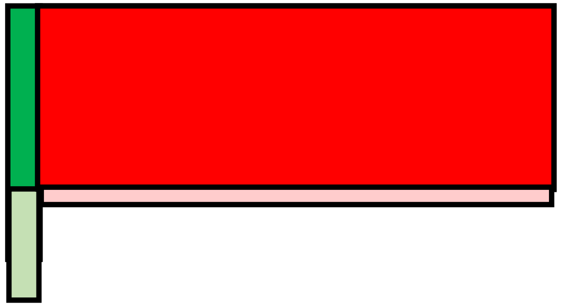
Agents will retain $\approx 50\%$ of the stories that they encounter from Pool O, and what they know is all desirable to know

→ having attractive stories “in the mix” may be essential to starting the social-filtering engine

As agents get stories from social sources, the veracity of neutral stories will then further increase



Pool O



retained inflow from Pool O
[similar to]
Pool S at threshold time $\hat{t}$

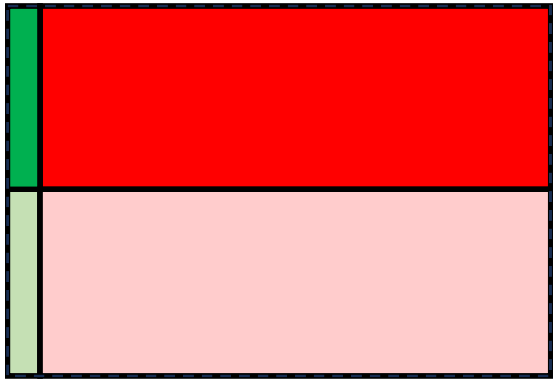
Example #2: Attractive Stories Stymie Social Filtering

Suppose $p_0 \approx 0, \hat{p} = \frac{1}{2}$, and $\pi_F + b = 0$

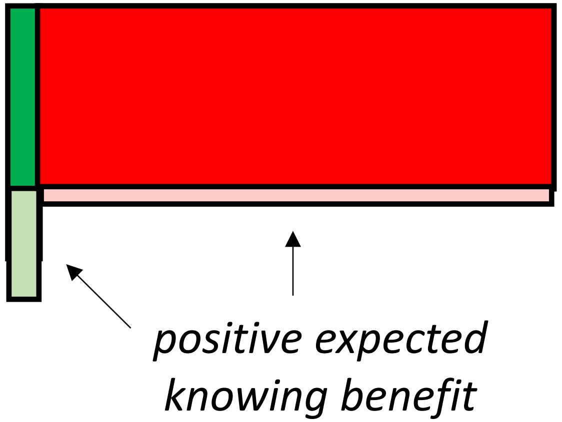
The attractive stories in social circulation are "junk" that crowds out the more-valuable neutral stories in circulation.

Even though knowledge prevalence is higher, the benefit of social interaction is lower than if all stories had been neutral

→ adding attractive stories to the mix may cause the Social Phase never to take off!



Pool O



Example #2: Attractive Stories Stymie Social Filtering

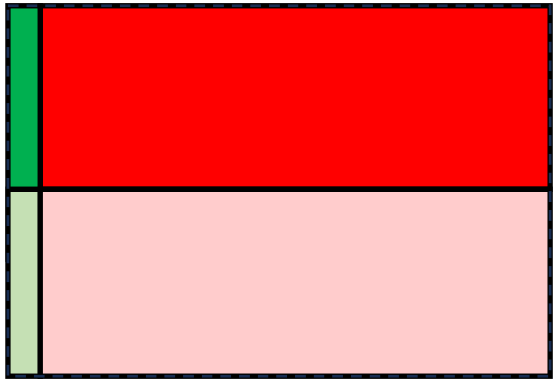
Suppose $p_0 \approx 0, \hat{p} = \frac{1}{2}$, and $\pi_F + b = 0$

The attractive stories in social circulation are "junk" that crowds out the more-valuable neutral stories in circulation.

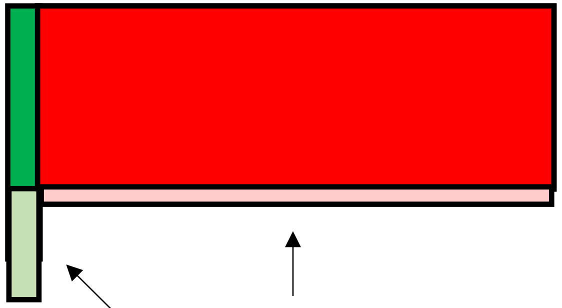
Even though knowledge prevalence is higher, the benefit of social interaction is lower than if all stories had been neutral

→ adding attractive stories to the mix may cause the Social Phase never to take off!

Remainder of talk: return to the No-Forgetting Assumption, which ensures that the Isolated Phase eventually ends.



Pool O



Plan for the Talk

Benchmark Model

Baseline Analysis

Attractive-Story Extension

- modeling “attractiveness”
- impact on attention dynamics
- **impact on veracity dynamics**
- equilibrium welfare implications

Partisan-Agent Extension

Story-Pool Dynamics: New Notation

$M_{\omega\chi}(t) =$ mass of stories with truth-value ω and attractiveness χ

$p_{S\chi}(t) \equiv \frac{M_{T\chi}(t)}{M_{T\chi}(t)+M_{F\chi}(t)} =$ social veracity of attractiveness-type-χ stories

$q(t) \equiv \frac{M_{T+}(t)+M_{F+}(t)}{M_{T+}(t)+M_{F+}(t)+M_{T0}(t)+M_{F0}(t)} =$ share of attractive stories in Pool S

$$M'_{T+}(t) = \underbrace{\lambda_O p_O q_O}_{\textit{inflow of true-attractive stories from Pool O}} + \underbrace{\lambda_S K(t) q(t) p_{S+}(t)}_{\textit{inflow of true-attractive stories from Pool S}} - \underbrace{\gamma M_{T+}(t)}_{\textit{outflow = zero}}$$

Note: There are no "σ" terms in this equation because attractive-story retention is 100%.

Note: Attention levels λ_O, λ_S are constant during the Social Phase.

Social-Veracity Dynamics: Attractive Stories

$$M'_{T+}(t) = \lambda_O q_O p_O + \lambda_S K(t) q(t) p_{S+}(t)$$

$$\frac{M'_{T+}(t)}{M_{T+}(t)} = \frac{1}{\Sigma M(t)} \left(\frac{\lambda_O q_O p_O}{q(t) p_{S+}(t)} + \lambda_S K(t) \right)$$

$$M'_{F+}(t) = \lambda_O q_O (1 - p_O) + \lambda_S K(t) q(t) (1 - p_{S+}(t))$$

$$\frac{M'_{F+}(t)}{M_{F+}(t)} = \frac{1}{\Sigma M(t)} \left(\frac{\lambda_O q_O (1 - p_O)}{q(t)(1 - p_{S+}(t))} + \lambda_S K(t) \right)$$

Social-Veracity Dynamics: Attractive Stories

$$\frac{M'_{T+}(t)}{M_{T+}(t)} - \frac{M'_{F+}(t)}{M_{F+}(t)} = \frac{\lambda_O q_O}{\Sigma M(t)q(t)}\left(\frac{p_O}{p_{S+}(t)} - \frac{1-p_O}{1-p_{S+}(t)}\right)$$

which equals zero at $p_{S+}(t) = p_O$.

→ $\boxed{p^*_{S+}(t) = p_O \text{ at all times } t}$

Social-Veracity Dynamics: Neutral Stories

$$M'_{T0}(t) = \lambda_O(1 - q_O)p_O\sigma_{OT} \; + \; \lambda_S K(t)\big(1 - q(t)\big)p_{S0}(t)\sigma_T(p_{S0}(t))$$

$$\frac{M'_{T0}(t)}{M_{T0}(t)} = \frac{1}{\Sigma M(t)}\left(\frac{\lambda_O(1 - q_O)p_O\sigma_{OT}}{(1 - q(t))p_{S0}(t)} + \lambda_S K(t)\sigma_T(p_{S0}(t))\right)$$

$$M'_{F0}(t) = \lambda_O(1 - q_O)(1 - p_O)\sigma_{OF} \; + \; \lambda_S K(t)\big(1 - q(t)\big)(1 - p_{S0}(t))\sigma_F(p_{S0}(t))$$

$$\frac{M'_{F0}(t)}{M_{F0}(t)} = \frac{1}{\Sigma M(t)}\left(\frac{\lambda_O(1 - q_O)(1 - p_O)\sigma_{OF}}{(1 - q(t))(1 - p_{S0}(t))} + \lambda_S K(t)\sigma_F(p_{S0}(t))\right)$$

Social-Veracity Dynamics: Neutral Stories

$$\frac{M'_{T0}(t)}{M_{T0}(t)} - \frac{M'_{F0}(t)}{M_{F0}(t)}$$

$$= \frac{\lambda_O(1-q_O)}{\Sigma M(t)(1-q(t))}\left(\frac{p_O\sigma_{OT}}{p_{S0}(t)} - \frac{(1-p_O)\sigma_{OF}}{1-p_{S0}(t)}\right) + \frac{1}{\Sigma M(t)}\lambda_S K(t)\Delta\sigma_T(p_{S0}(t))$$

$$= \frac{1-q_O}{1-q(t)} IN\big(t; p_{S0}(t)\big) + SF\big(t; p_{S0}(t)\big) < IN\big(t; p_{S0}(t)\big) + SF\big(t; p_{S0}(t)\big)$$

→ neutral-story social veracity, $p^*_{S0}(t)$, (i) **rises more slowly** and (ii) **converges to a lower long-run limit** than in the benchmark model

Key Finding: Attractive Stories

Theorem 2. In the attractive-story extension:

(i) $q^*(t) > q_O$ for all t.

(ii) $p^*_{S+}(t) = p_O$ for all t.

(iii) $p_S^{*\infty} > p_{S0}^{*\infty} > p_O$

Plan for the Talk

Benchmark Model

Baseline Analysis

Attractive-Story Extension

- modeling “attractiveness”
- impact on attention dynamics
- impact on veracity dynamics
- **equilibrium welfare implications**

Partisan-Agent Extension

Welfare Implications

If the benefit b is large enough, then obviously agents will be better off once attractive stories enter the info ecosystem.

Otherwise, adding attractive stories to the mix can lower long-run equilibrium agent welfare, for two reasons:

- *attractive stories are not subjected to social filtering* → the attractive stories in circulation have lower veracity than they otherwise would
- *neutral stories are "crowded out" by attractive ones* → high-veracity neutral stories that otherwise would have been told and told again have shorter transmission chains and reach fewer agents

Example: Attractive Stories Reduce Welfare

– $b = -\pi_F + \epsilon$, so that false-attractive stories generate ≈ zero benefit
– $p_O \approx 0$, so that almost all stories are false
– $q_O = \frac{1}{2}$, so that neutral and attractive stories are inherently equally common

Neutral Stories: The neutral-story inflow from Pool O is tiny but has relatively high veracity, equal to $p_{in}(p_O) > \hat{p}$.

Attractive Stories: The attractive-story inflow from Pool O has veracity p_O and, in volume, dwarfs the neutral-story inflow.

→ almost all stories that people know are attractive, i.e., $q^*(t) \approx 1$ for all t

→ stories encountered socially generate ≈ zero expected benefit

Plan for the Talk

Benchmark Model

Baseline Analysis

Attractive-Story Extension

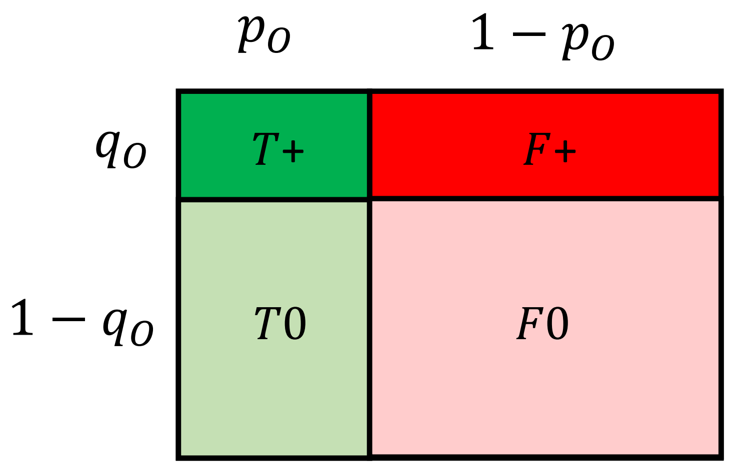
Partisan-Agent Extension

- modeling agent heterogeneity
- can “echo chambers” persist in the long run?

Model: Stories & Agents

Stories

- each original story is (i) *true* or *false* ($\omega = T, F$) and (ii) *attractive* or *neutral* ($\chi = +, 0$), as in the attractive-story extension

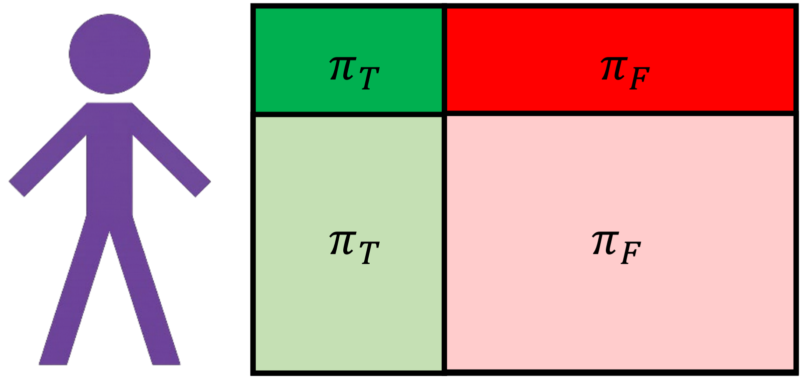


original stories
(Pool O)

Model: Stories & Agents

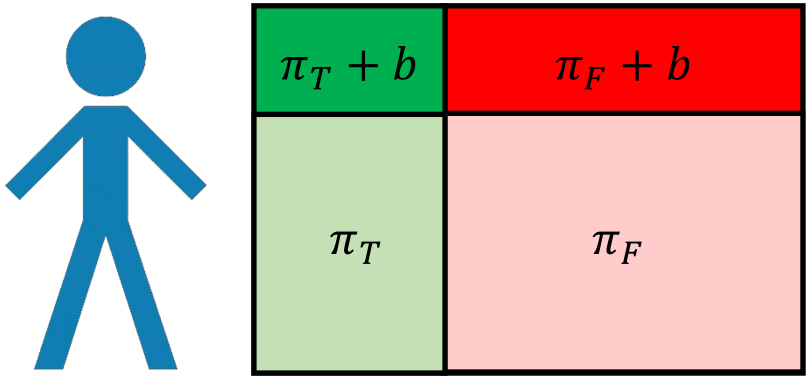
Stories

- each original story is (i) *true* or *false* ($\omega = T, F$) and (ii) *attractive* or *neutral* ($\chi = +, 0$), as in the attractive-story extension



Agents

- two types: non-partisan (NP) and partisan (P)
- NP agents get knowing payoffs as in the benchmark model
- P agents get knowing payoffs as in the attractive-story extension

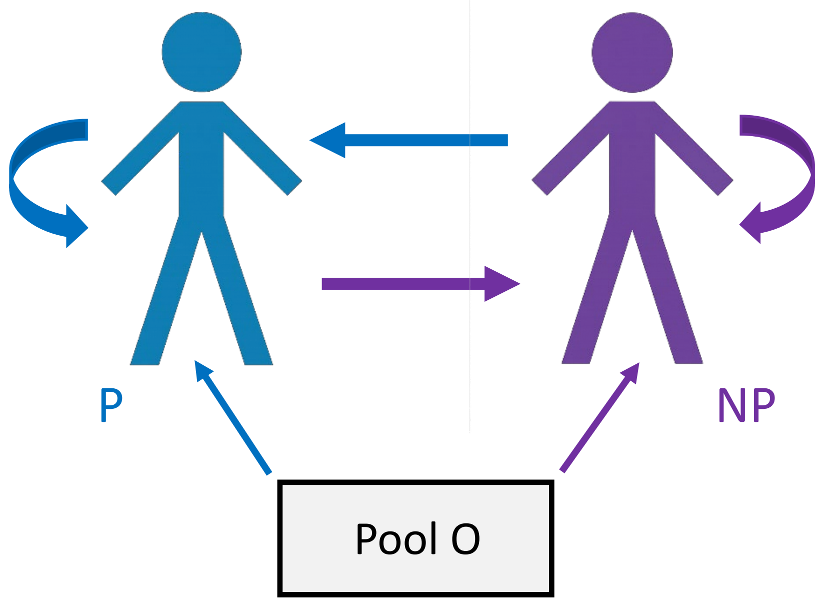


Model: Attention

Agents can allocate their attention budget any way they like across

- original sources
- social interaction with own type
- social interaction with other type

A *"type-θ echo chamber"* is a situation in which type-θ agents interact among themselves but never with the other type



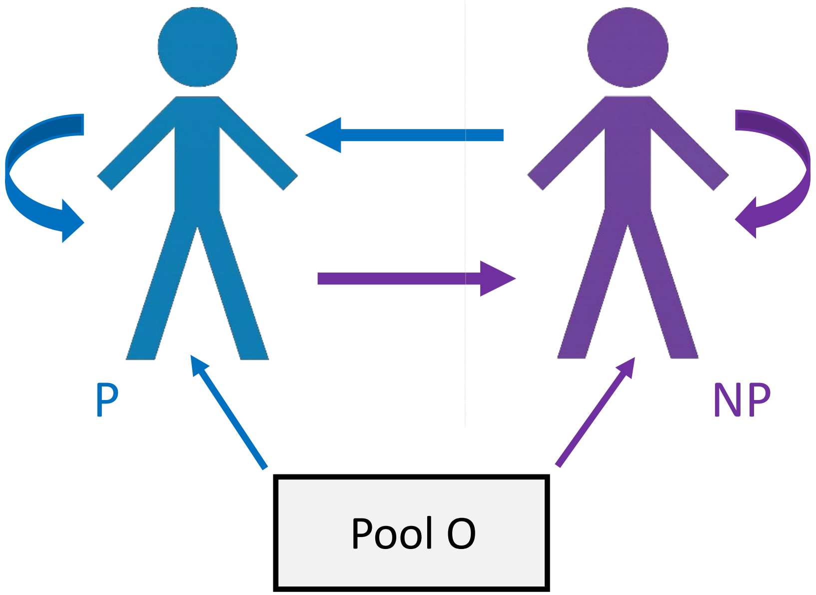
Non-Partisan Agents: Higher Veracity

NP agents (“NPs”) pay full attention to the source with the highest overall social veracity

AND

NPs are selective in the attractive stories they choose to retain, while Ps are not

→ in the long run, the veracity of stories known by NPs must strictly exceed the veracity of stories known by Ps



Non-Partisan Agents: Echo Chamber

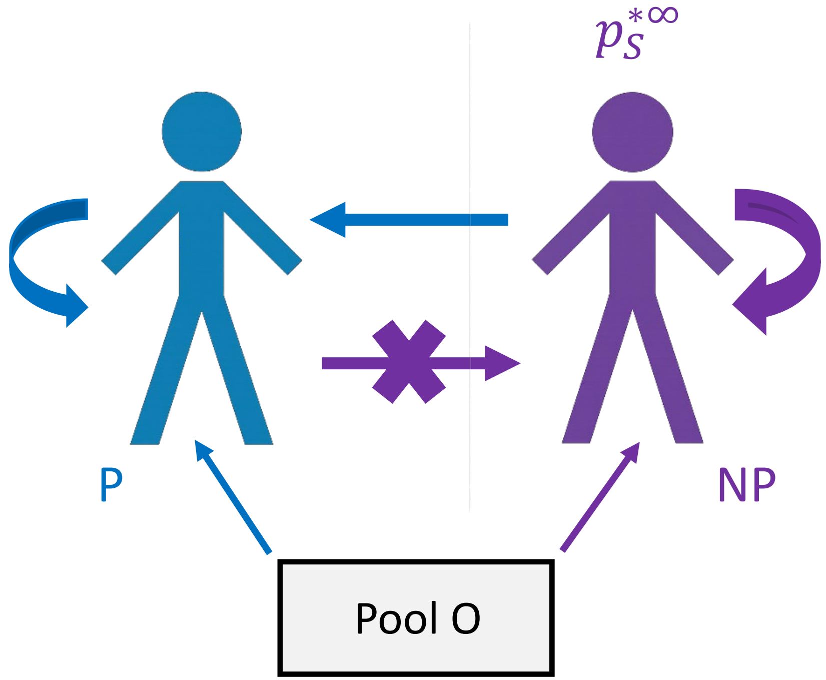
NP agents ("NPs") pay full attention to the source with the highest overall social veracity

AND

NPs are selective in the attractive stories they choose to retain, while Ps are not

→ in equilibrium, an NP echo chamber must eventually emerge and persist forever

→ $\boxed{p_{S+}^{NP*\infty} = p_{S0}^{NP*\infty} = p_S^{*\infty}}$ and $\boxed{q^{*\infty} = q_O}$



Partisan Agents: A Partisan Echo Chamber?

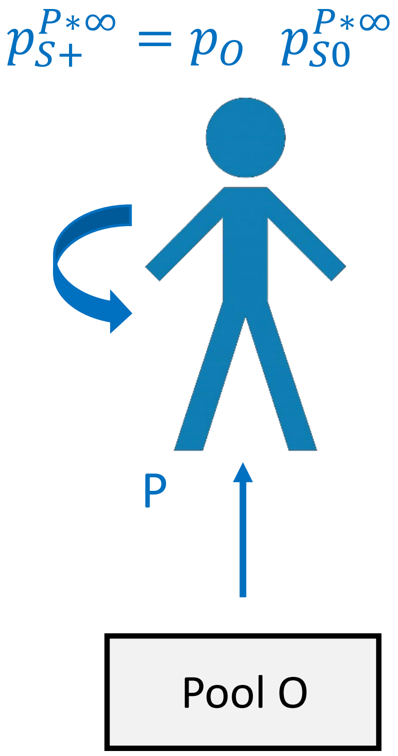
Suppose that Ps also only interact with Ps.

Long-run veracity and prevalence in Pool P must then be as in the attractive-story extension. Namely:

- $p_{S+}^{P*\infty} = p_O$
- $q^{P*\infty} > q_O$
- $p_{S0}^{P*\infty} < p_S^{*\infty}$

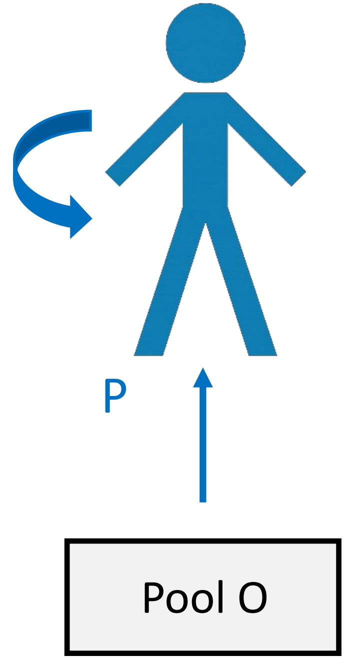
There is a clear tradeoff for Ps when deciding whether to "follow" Ps vs NPs

- stories known by NPs are *more true*
- stories known by Ps are *more attractive*



Partisan Agents: A Partisan Echo Chamber?

$p_{S+}^{P*\infty} = p_O \quad p_{S0}^{P*\infty}$



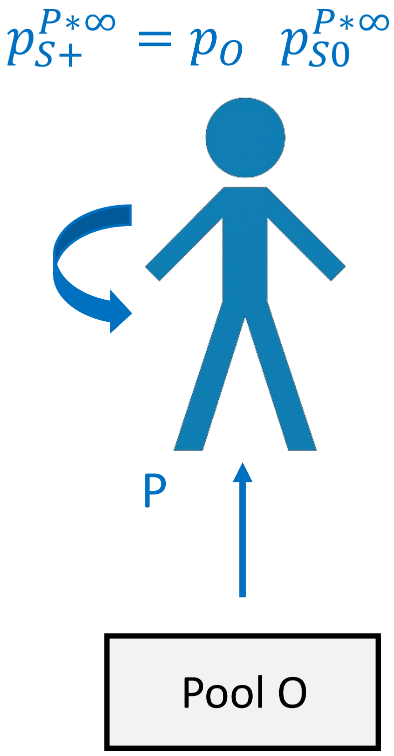
For a P agent, staying in the partisan echo chamber yields per-story expected payoff

$$q^{P*\infty}\underbrace{(b + p_O\pi_T + (1 - p_O)\pi_F)}_{\textit{P agents always retain attractive stories}} + (1 - q^{P*\infty})\underbrace{E\Pi\left(p_{S0}^{P*\infty}\right)}_{\textit{for neutral stories, P agents follow the same retention strategy and get the same payoffs as NPs}}$$

By contrast, stories heard from NPs have expected payoff

$$q_O(b + p_S^{*\infty}\pi_T + (1 - p_S^{*\infty})\pi_F) + (1 - q_O)E\Pi(p_S^{*\infty})$$

Partisan Agents: A Partisan Echo Chamber?



Consider the case with $p_O \approx 0$, i.e., ~ all stories are false.

Staying in the partisan echo chamber yields per-story expected payoff $b + \pi_F$.

Deviating to “listen” to NP agents yields per-story payoff

$$\underbrace{q_O(b + p_S^{*\infty}\pi_T + (1 - p_S^{*\infty})\pi_F)}_{\text{since } p_S^{*\infty} > \hat{p} \gg 0\text{, P agents prefer to hear attractive stories from NP sources}} + (1 - q_O)\underbrace{E\Pi(p_S^{*\infty})}_{\text{P agents also benefit from the neutral stories they hear from NP sources}}$$

There is a unique $\hat{b} > -\pi_F$ such that P agents prefer not to leave the echo chamber if and only if $b > \hat{b}$.

Summary and Key Takeaways

This paper introduced a simple and tractable model of *the evolution of collective knowledge* when people are exposed to stories socially.

Sociality drives a virtuous circle, in which social filtering of information makes social interaction more valuable.

Socially-encountered stories with “viral” features are (endogenously) less likely to be true.

When original-story veracity is low, “partisan echo chambers” may emerge but only if partisan agents’ benefit from knowing false partisan-aligned content is sufficiently high.

Exciting Future Directions: Info Ecology

Bits of information do not exist in isolation. They interact ecologically within **infobiomes**, much as bacteria interact within microbiomes.

Specific infobiome topics that especially interest me:

- competing ideas
  - see Dasaratha and He ("Learning from Viral Content")
- legibility & enabling knowledge
- perspectives & ideology