Coupled Contagions, Agent_Zero, and Generative Social Science

Joshua M. Epstein

Professor of Epidemiology

Director, the NYU Agent-Based Modeling Lab

June 4, 2020

Bio Dynamics Days 2020

Le Havre Normandie University and the Courant Institute, NYU

Affiliations

- Professor of Epidemiology NYU
- Director, the NYU ABM_Lab
  - Several Courant members (Benjamin, Lai-Sang, others)
- Affiliated Appointments to Courant and the School of Arts and Sciences

Three Parts

1. Coupled Contagion Dynamics of Fear and Disease

2. Agent_Zero: a neurocognitive foundation for generative social science

3. Extensions/applications/collaborative opportunities

    --Large-Scale Models

    --Inverse Generative Social Science w Machine Learning

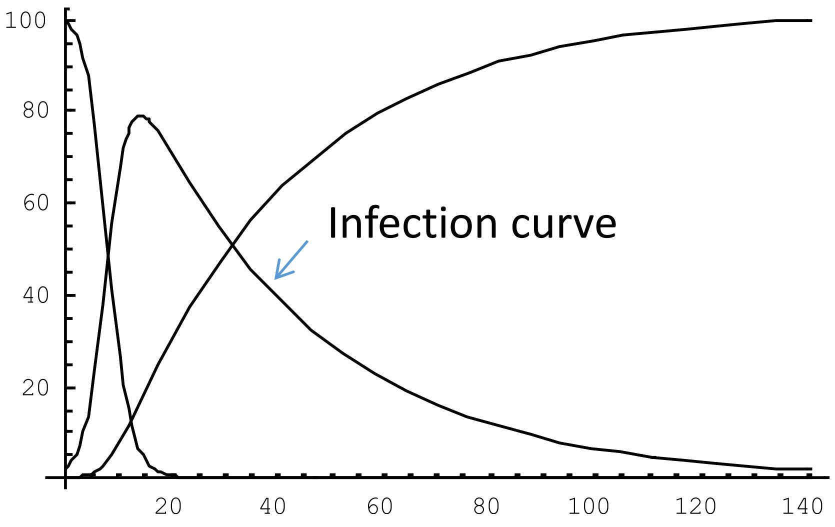
To begin, the classic model…

The Classic Kermack-McKendrick (1927) SIR Model* Illuminated Thresholds and Herd Immunity

Susceptible Growth Rate: $dS/dt = -\beta SI$

Infective Growth Rate: $dI/dt = \beta SI - \Upsilon I$

Removed Growth Rate: $dR/dt = \Upsilon I$



SI → Perfect Mixing.

* See Hoppensteadt, F. C., & Peskin, C. S. (2012). *Modeling and simulation in medicine and the life sciences* (Vol. 10). Springer

Gave Counterintuitive Insights

Showed that epidemics are Threshold Phenomena. How?

"Epidemic" means $\dfrac{dI}{dt} > 0,$

which is to say $\beta SI - \gamma I > 0.$

Takes off if $S > \dfrac{\gamma}{\beta}$

Fizzles if $S < \dfrac{\gamma}{\beta}$

For "Herd Immunity," just vaccinate until remaining Susceptible pool is sub-threshold. The classic result for homogeneous and well-mixed populations. Also, overly deadly bugs aren't very good at sustaining epidemics.

Recast in Terms of Famous R_0

$$R_0 \equiv \frac{\beta S_0}{\gamma}$$

Interpretation: net secondary infections for one index

$$R_0 > 1 \Rightarrow \text{epidemic}$$
$$R_0 < 1 \Rightarrow \text{none}$$

In general, R_0 is the spectral radius of the so-called Next-Generation Matrix.

$\mathrm{R}_0 > 1 \rightarrow$ instability of the zero equilibrium.

Pretty Formula for Herd Immunity by Vaccination (Perfect)

Vaccinees are subtracted from S $$S_0 - v < \frac{\gamma}{\beta}$$

Equivalently: $$v > 1 - \frac{1}{R_0}$$

Pretty darn neat!

1918 pandemic flu R_0 =2, so vaccination of ½ would have sufficed. Peak Ebola was also around 2.0. COVID-19 varies but has also reached this.

All in all…

- Triumph of elegant analytic modeling.
- Picasso, “Art is a Lie that helps us see the truth.”
- So it is for all the best models.
- Gave important qualitative insights, and even works well empirically when perfect mixing applies…

But there’s only one problem…

But there's only one problem…

- Perfect mixing: Would you keep milling around town if you knew there were a plague underway?
- Doubtful…

"The plague was nothing; fear of the plague was much more formidable."

Henri Poincare

Classical models do not include any behavioral adaptation…Why not?

- Most modelers would say, “Human behavior’s just too hard, so we leave it out.”
- No Dice: You aren’t leaving it out!
- You are making a strong behavioral assumption (invariant contact pattern).
- So, you are including behavior one way of the other…just badly.
- Can we do a little better?
- Here are a couple simple approaches…all extensible with students and faculty (several publications possible).

Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations (Epstein et al, 2008 Plos_ONE)

- Two interacting contagion processes: one of disease one of *fear about* the disease.
- Individuals contract disease only through contact with the disease-infected (the sick).
- Individuals contract fear through contact with the disease-infected (the sick), the fear-infected (the scared), and those infected with both fear and disease (the sick and scared).
- Scared individuals--whether sick or not--withdraw from circulation with some probability, which affects the course of the disease epidemic proper.
- If individuals recover from fear and return to circulation, the disease dynamics become rich, and include multiple waves of infection, such as occurred in the 1918 flu.
- Recent work on this using Twitter Data (Broniatowski, et al, 2016).

Two Formulations

- Differential Equations
- Agent-Based Computational Model

Possible states

S: Susceptible to pathogen and fear
I_F: Infected with fear only
I_P: Infected with pathogen only
I_{PF}: Infected with pathogen and fear
R_F: Removed from circulation due to fear
R_{PF}: Removed from circulation due to fear and infected with pathogen
R: Recovered from pathogen and immune to fear

Transmission probabilities

	Get scared (α)	Not get scared
Get Sick (β)	$\alpha\ \beta$	$(1-\alpha)\ \beta$
Not get sick	$\alpha\ (1-\beta)$	$(1-\alpha)\ (1-\beta)$

Parameters governing Removal and Return

λ_1: Rate of removal to self-isolation of those infected with fear only
λ_2: Rate of recovery from infection with pathogen
λ_3: Rate of removal to self-isolation of those infected with fear and pathogen
H: Rate of recovery from fear and return to circulation

Possible states

S: Susceptible to pathogen and fear
I_F: Infected with fear only
I_P: Infected with pathogen only
I_{PF}: Infected with pathogen and fear
R_F: Removed from circulation due to fear
R_{PF}: Removed from circulation due to fear and infected with pathogen
R: Recovered from pathogen and immune to fear

Transmission probabilities

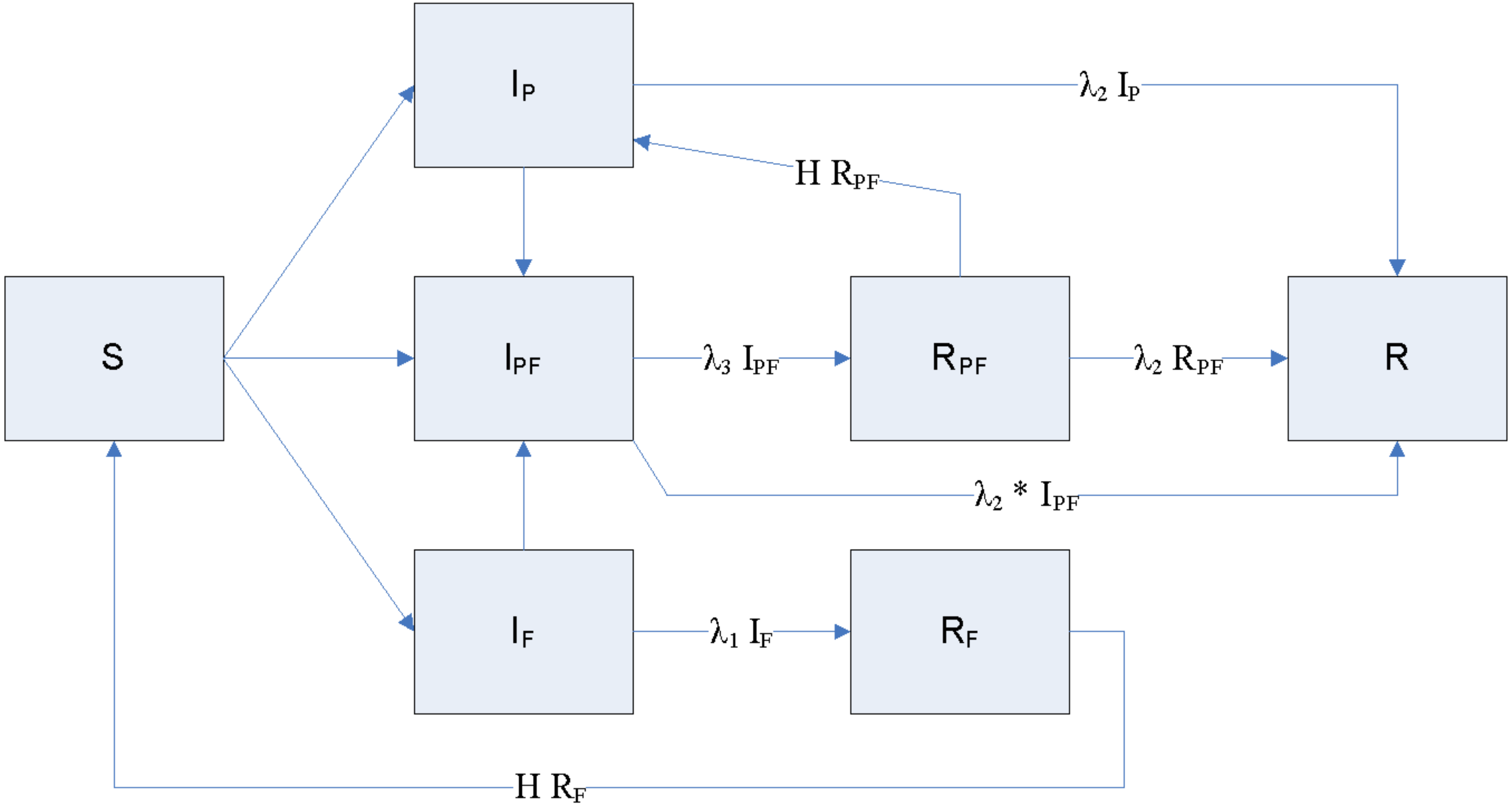
	Get scared (α)	Not get scared
Get Sick (β)	$\alpha\ \beta$	$(1-\alpha)\ \beta$
Not get sick	$\alpha\ (1-\beta)$	$(1-\alpha)\ (1-\beta)$

Salem witches…

Parameters governing Removal and Return

λ_1: Rate of removal to self-isolation of those infected with fear only
λ_2: Rate of recovery from infection with pathogen
λ_3: Rate of removal to self-isolation of those infected with fear and pathogen
H: Rate of recovery from fear and return to circulation

Agincourt state transition chart



Classical SIR Differential Equations Formulation

$$\frac{dS}{dt} = -\beta(1-\alpha)SI_P - (1-\beta)\alpha SI_P - \beta\alpha SI_P - \alpha SI_F - \beta(1-\alpha)SI_{PF}$$
$$- (1-\beta)\alpha SI_{PF} - \beta\alpha SI_{PF} + HR_F$$

$$\frac{dI_F}{dt} = (1-\beta)\alpha SI_P + \alpha SI_F + (1-\beta)\alpha SI_{PF} - \beta I_F I_P - \beta I_F I_{PF} - \lambda_1 I_F$$

$$\frac{dI_P}{dt} = \beta(1-\alpha)SI_P + \beta(1-\alpha)SI_{PF} - \alpha I_P I_P - \alpha I_P I_F - \alpha I_P I_{PF} - \lambda_2 I_P + HR_{PF}$$

$$\frac{dI_{PF}}{dt} = \beta\alpha SI_P + \beta\alpha SI_{PF} + \beta I_F I_P + \beta I_F I_{PF} + \alpha I_P I_P + \alpha I_P I_F + \alpha I_P I_{PF} - \lambda_2 I_{PF} - \lambda_3 I_{PF}$$

$$\frac{dR_F}{dt} = \lambda_1 I_{PF} - HR_F$$

$$\frac{dR_{PF}}{dt} = \lambda_3 I_{PF} - \lambda_2 R_{PF} - HR_{PF}$$

$$\frac{dR}{dt} = \lambda_2 I_P + \lambda_2 I_{PF} + \lambda_2 R_{PF}$$

Subsumes Classical Models

With β=0, SIR for fear

With α=0, SIR for pathogen

Removals and re-entry to S

$$\frac{dS}{dt} = -\beta(1-\alpha)SI_P - (1-\beta)\alpha SI_P - \beta\alpha SI_P - \alpha SI_F - \beta(1-\alpha)SI_{PF}$$
$$- (1-\beta)\alpha SI_{PF} - \beta\alpha SI_{PF} + HR_F$$

$$\frac{dI_F}{dt} = (1-\beta)\alpha SI_P + \alpha SI_F + (1-\beta)\alpha SI_{PF} - \beta I_F I_P - \beta I_F I_{PF} - \lambda_1 I_F$$

$$\frac{dI_P}{dt} = \beta(1-\alpha)SI_P + \beta(1-\alpha)SI_{PF} - \alpha I_P I_P - \alpha I_P I_F - \alpha I_P I_{PF} - \lambda_2 I_P + HR_{PF}$$

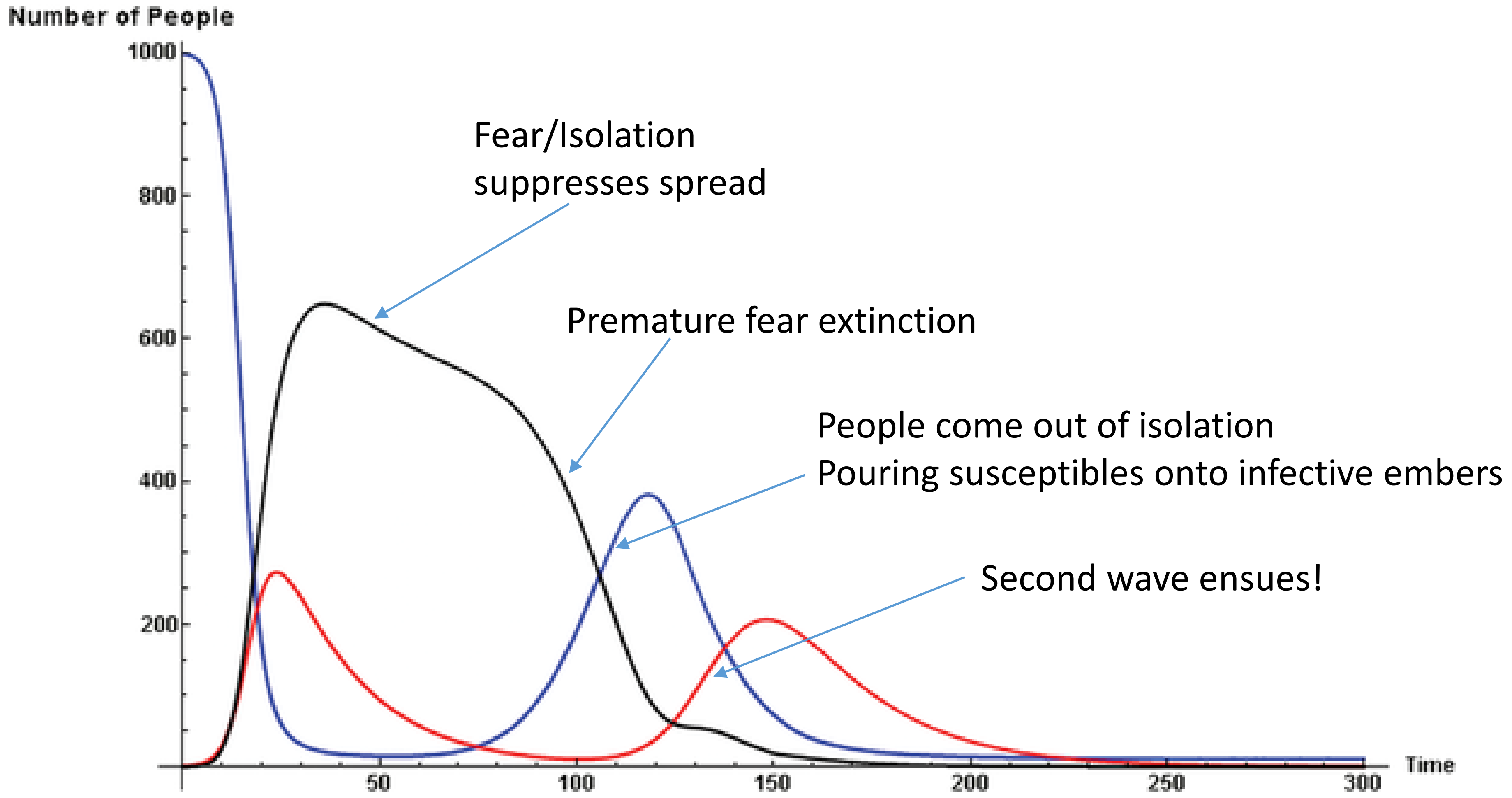
$$\frac{dI_{PF}}{dt} = \beta\alpha SI_P + \beta\alpha SI_{PF} + \beta I_F I_P + \beta I_F I_{PF} + \alpha I_P I_P + \alpha I_P I_F + \alpha I_P I_{PF} - \lambda_2 I_{PF} - \lambda_3 I_{PF}$$

$$\frac{dR_F}{dt} = \lambda_1 I_{PF} - HR_F$$

$$\frac{dR_{PF}}{dt} = \lambda_3 I_{PF} - \lambda_2 R_{PF} - HR_{PF}$$

$$\frac{dR}{dt} = \lambda_2 I_P + \lambda_2 I_{PF} + \lambda_2 R_{PF}$$

Figure 3. In the idealized run of figure 3, susceptible individuals (blue-curve) self-isolate (black curve) through fear as the infection of disease proper grows (red curve).



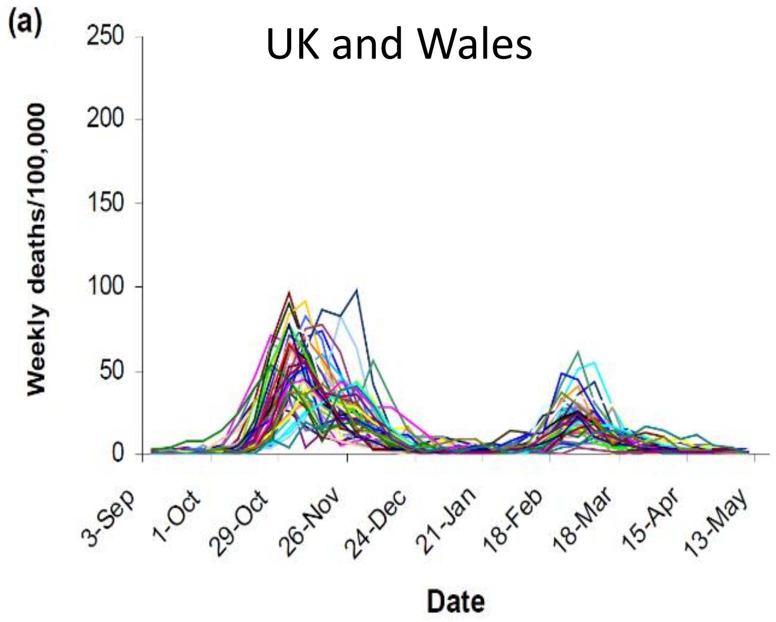
Epstein JM, Parker J, Cummings D, Hammond RA (2008) Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations. PLOS ONE 3(12): e3955. https://doi.org/10.1371/journal.pone.0003955
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0003955

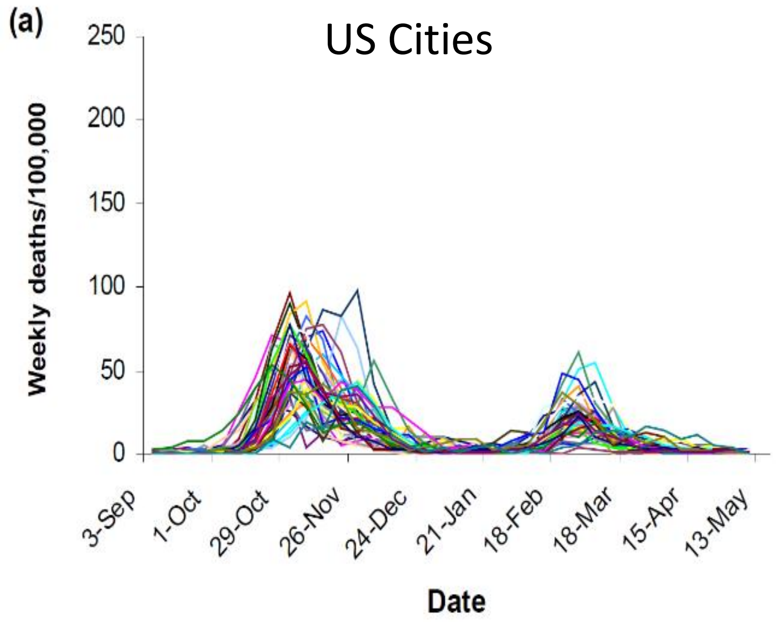
PLOS | ONE

Main Mystery of 1918 Spanish Flu (50m dead)

- Multiple waves of infection.
- Here is a behavioral mechanism

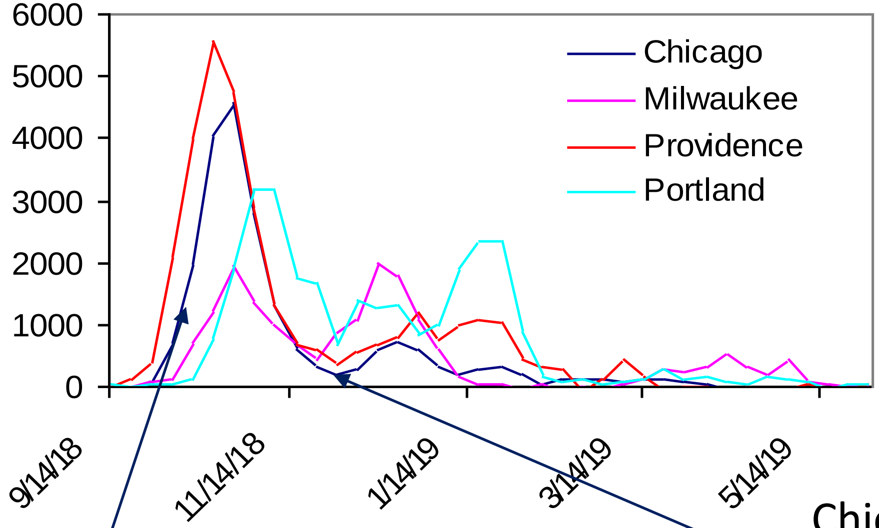
Very Crude Qualitative Agreement for Cities





Emboldened, I conducted a massive *big data* analysis for Chicago

Chicago



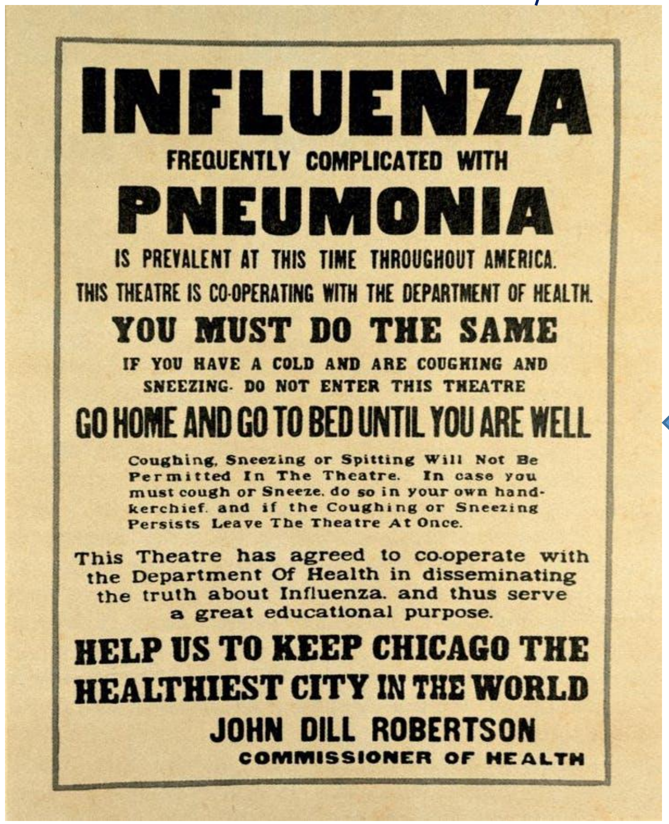
Chicago Tribune, November 10, 1918

ALL BANS OFF

Chicago Healthiest City in the World, Says Robertson.

If you want to dodge the "flu" and the "pneu" Chicago is the best place to be in. The epidemic here is over. Dr. John Dill Robertson, commissioner of health, in a letter to Lucius Teter, president of the Chicago Association of Commerce, yesterday declared that, so far as pneumonia and influenza are concerned, Chicago is the safest place in the United States.

"We are practically out of the woods," wrote Dr. Robertson. "All bans are off. In a few days I am sure I shall again be justified in stating that Chicago is the healthiest city in the world."



Adages…

- Einstein, Theory should be “as simple as possible, but no simpler”
- Epstein, Data should be “as big as necessary, but no bigger.”

For full analysis, see Bootsma and Ferguson (PNAS 2007 104 (18) 7588-7593) actually did detailed analyses for each city.

Fear vs. Bug

- Fear can dissipate before the infection does
- Fear can also transmit faster then the bug itself, even if α=β (why?)
  - More channels
  - Does not require physical contact
- Fear faster than bug if:

$$\alpha > \frac{\beta\lambda_1(\lambda_2+\lambda_3)}{(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)-\beta\lambda_1\lambda_3}$$

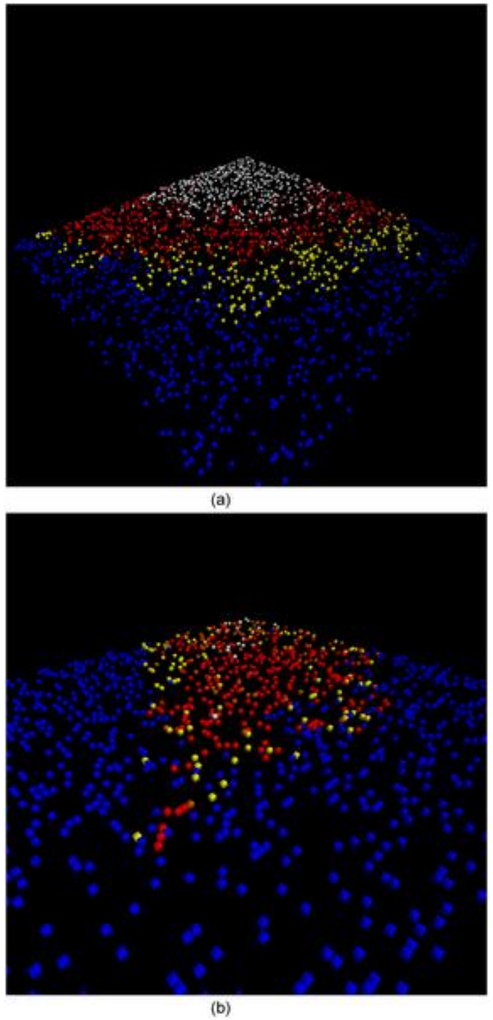
Fear can stimulate social flight…

- Even when there’s no disease
  - Surat India 1994, 350k fled pneumonic plague. World Health Organization. Zero cases.
- Flight is double-edged, where there is disease…

Agent-Based Computational Model (reaction-diffusion model also fertile, as in JD Murray's Black Plague model)

- Random movement on torus (200 x 200)
- Agent population (8000)
- I(0)=1
- Moore neighbor interactions
- Color code:
  - S_{BF} = blue
  - I_F = yellow
  - I_B = orange
  - I_{BF} = red
  - R_{BF} = light gray
  - R_F = white
  - R_B = dark gray (barely visible, by design)

Figure 9. (A&B): Screenshots from the agent-based simulation model without and with flight.



Epstein JM, Parker J, Cummings D, Hammond RA (2008) Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations. PLOS ONE 3(12): e3955. https://doi.org/10.1371/journal.pone.0003955
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0003955

Fear-Driven Flight…

- Even small levels of flight can amplify epidemic scale dramatically.
- Discussed some very rudimentary behaviors.
- Another wrinkle…we had bug and fear of bug…

Can also have fear of the control: a two-fear model

[1] Contagious Disease

[2] Contagious Fear of the Disease

[3] Contagious Fear of the Control Measure

The WHO just added vaccine refusal (fear-driven) to the top ten threats to global health. Resurgence of Polio, Pertussis, Measles…could undermine the COVID-19 vaccination program.

Two-Fear, Three Contagions Model

- C1: Measles proper (highly contagious $R_0 \approx 12$)
- C2: Fear of measles (legitimate)
- C3: Fear of the measles vaccine (Illegitimate)
- Dynamics depend on the relationship between the fears.

Thinking of some disease threat, the simplest toy model is this. This is the wrong model, but it's the right idea!

T(t) some threat

V(t) some control

$\emptyset_T$ fear of the threat … constant for now

$\emptyset_V$ fear of the control … constant for now

[1] $$\frac{dV}{dt} = (\emptyset_T - \emptyset_V)T$$

[2] $$\frac{dT}{dt} = -\xi V.$$

[3] $$\frac{d^2T}{dt^2} = -\xi(\emptyset_T - \emptyset_V)T$$ Hooke's Law with spring constant $\xi(\emptyset_T - \emptyset_V)$

Dynamically

- $(\emptyset_T - \emptyset_V) > 0$ Center (Hooke′s Law)
- $(\emptyset_T - \emptyset_V) < 0$ Saddle (Explosive growth)

- Lots of extensions: control fatigue and seasonal forcing

With $\emptyset_T > \emptyset_V$ and control fatigue, spiral sink (like 1918 cities)

$$\frac{dV}{dt} = (\emptyset_T - \emptyset_V)T - aV$$

$$\frac{dT}{dt} = -\xi V$$

OR, the periodically forced undamped case…pandemic as resonance catastrophe!

$$\frac{d^2T}{dt^2} = -\xi(\emptyset_T - \emptyset_V)T + ASin(\omega t)$$

Now the same impulse, but the “right “ model

The Two-Fear Model

w/ Erez Hatna (NYU AMB_Lab) and Jennifer Crodelle (Courant)

Case 1: pathogen

Case 2: pathogen + fear of pathogen

Case 3: pathogen + fear of pathogen + vaccinations

Case 4: pathogen + fear of pathogen + vaccinations + fear of vaccinations

General Equations forthcoming arXiv

$$\frac{dS}{dt} = -\beta IS + \alpha_{fd}S_{fd}R_{nat} + \alpha_{fv}S_{fv}R_{vac} - \beta_{fd}S_{fd}S - \hat{\beta}_{fd}IS - \beta_{fv}S_{fv}S - \hat{\beta}_{fv}AS + \mu I$$
$$\frac{dS_{fd}}{dt} = -p\beta IS_{fd} - vS_{fd} - \alpha_{fd}S_{fd}R_{nat} + \beta_{fd}S_{fd}S + \hat{\beta}_{fd}IS$$
$$\frac{dS_{fv}}{dt} = -\beta IS_{fv} - \alpha_{fv}S_{fv}R_{vac} + \beta_{fv}S_{fv}S + \hat{\beta}_{fv}AS$$
$$\frac{dI}{dt} = \beta IS + p\beta IS_{fd} + \beta IS_{fv} - \gamma I - \mu I$$
$$\frac{dR_{nat}}{dt} = \gamma I$$
$$\frac{dR_{vac}}{dt} = (1-\sigma)vS_{fd} + \alpha_{adv}A$$
$$\frac{dA}{dt} = \sigma vS_{fd} - \alpha_{adv}A$$
$$\frac{dv}{dt} = \eta(S_{fd} - S_{fv})v(\epsilon - v),$$

SIR for disease

Fear-of-disease

$$\frac{dS}{dt} = -\beta IS + \alpha_{fd}S_{fd}R_{nat} + \alpha_{fv}S_{fv}R_{vac} - \beta_{fd}S_{fd}S - \hat{\beta}_{fd}IS - \beta_{fv}S_{fv}S - \hat{\beta}_{fv}AS + \mu I$$

Fear of vaccine

$$\frac{dS_{fd}}{dt} = -p\beta IS_{fd} - vS_{fd} - \alpha_{fd}S_{fd}R_{nat} + \beta_{fd}S_{fd}S + \hat{\beta}_{fd}IS$$

$$\frac{dS_{fv}}{dt} = -\beta IS_{fv} - \alpha_{fv}S_{fv}R_{vac} + \beta_{fv}S_{fv}S + \hat{\beta}_{fv}AS$$

$$\frac{dI}{dt} = \beta IS + p\beta IS_{fd} + \beta IS_{fv} - \gamma I - \mu I$$

$$\frac{dR_{nat}}{dt} = \gamma I$$

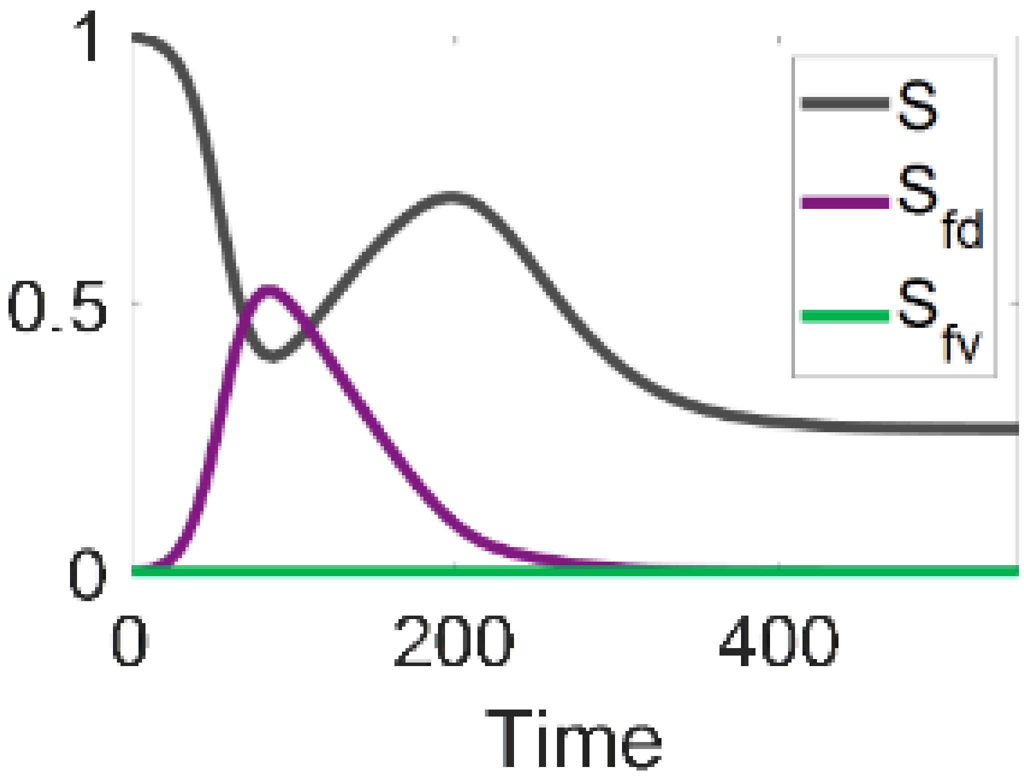
$$\frac{dR_{vac}}{dt} = (1-\sigma)vS_{fd} + \alpha_{adv}A$$

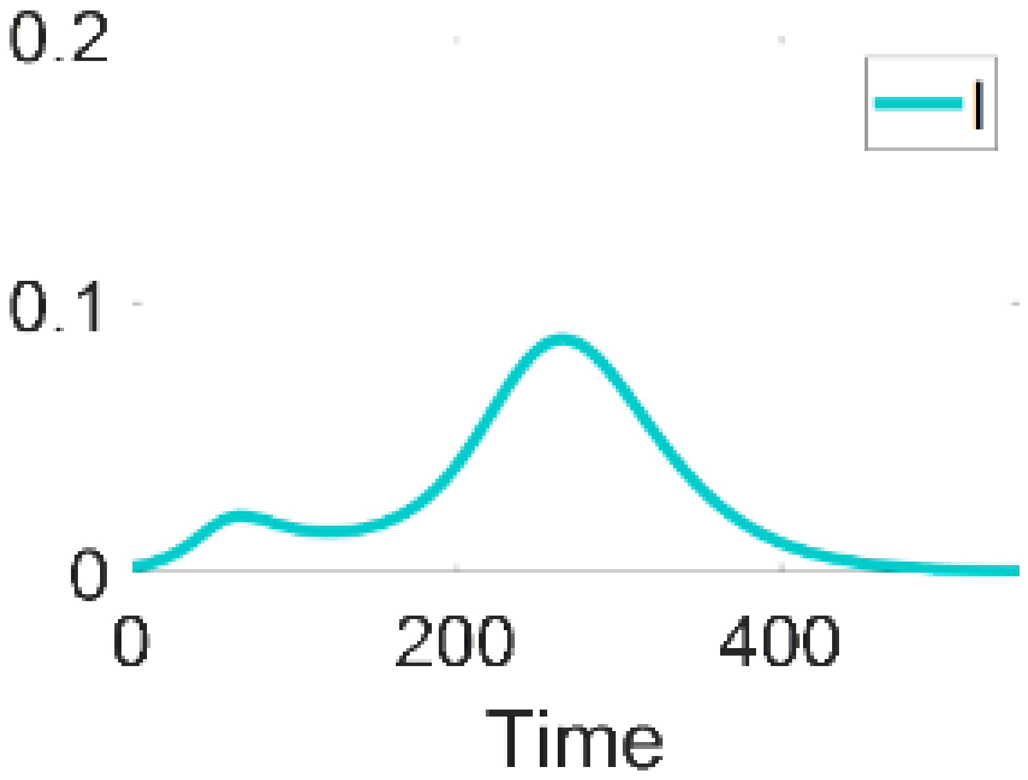
Difference between fears governs vaccine uptake

$$\frac{dA}{dt} = \sigma vS_{fd} - \alpha_{adv}A$$

$$\frac{dv}{dt} = \eta(S_{fd} - S_{fv})v(\epsilon - v),$$

One Run with no Vaccine (a la 1918) Second Wave Worse Than The First





Smallpox Vaccine Cycles

In her wonderful social history of smallpox, the *Speckled Monster,* Jennifer Lee Carrell recounts, “In London, inoculation’s popularity waxed and waned through the 1730s, with the force of the disease: in bad years, people flocked to be inoculated; in light years, the practice shrank. Inoculation was a security—the *only* security—to cling to within the terror of an epidemic; in times of good health, however, it looked like a foolish flirtation with danger.”

No Unified Behavioral Model

- Typically two control mechanisms
- Vaccination
- Isolation/Social Distancing
- Both driven by behaviors
- There is no model unifying these.
- With quarantine lags (as in Lai-Sang’s model), networks, spatial structure, it would be rich and important
- OK, so those are some thoughts on how to get coupled fear-driven behaviors into mathematical epidemiology

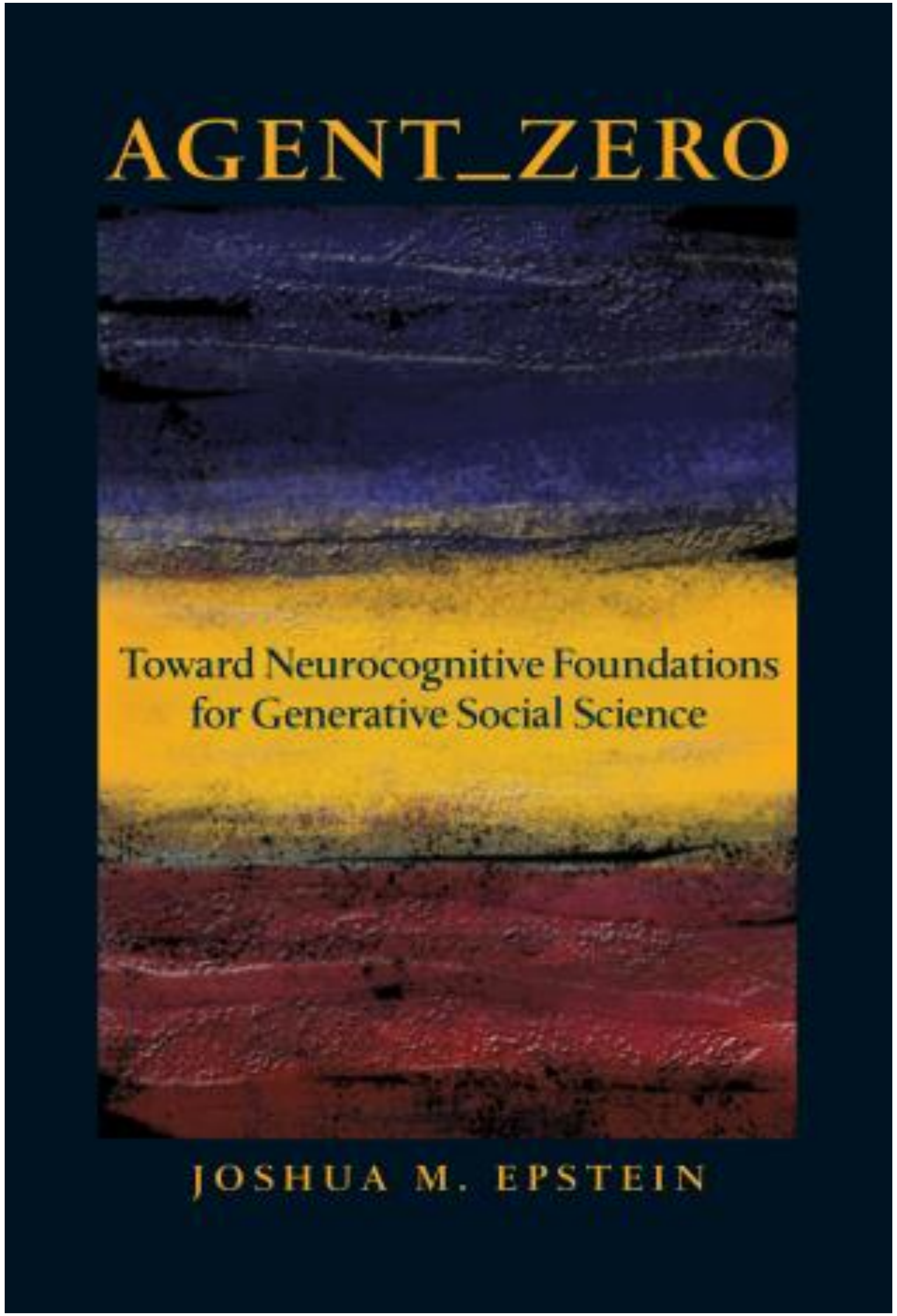
Taking "Fear" Seriously…

- We've thrown this term around…all very nice.
- But can treat it more seriously?
- How fear happens, how it evaporates, any neural correlates.
- I know Benjamin and others are interested in bringing mathematical neuroscience into all this.
- One humble provisional attempt to do that is…

Agent_Zero

Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science
Princeton University Press 2013

Funded by an NIH Director's Pioneer Award



Third in a Trilogy on Agent-Based Modeling

- Epstein and Axtell, *Growing Artificial Societies: Social Science from the Bottom Up* (MIT Press, 1996).
  - Exploratory
  - Immature Epistemology
- Epstein, *Generative Social Science: Studies in Agent-Based Computational Modeling* (Princeton Press, 2006).
  - Explanatory: Artificial Anasazi, Epidemics, Civil Violence, Classes, Retirement, Organizations.
  - Mature Epistemology
- Epstein, *Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science* (Princeton Press, 2013)
  - Cognitively plausible agent as foundation for generative explanations

Generative explanation*

- To explain a social regularity
- Demonstrate how it could emerge on time scales of interest to humans in a population of cognitively plausible agents
- Does the micro-specification *m* generate the macroscopic *explanandum x*
- If so, *m* is a generative explanatory candidate.
- Motto (Epstein, 1999) is negative : If you didn't grow it, you didn't explain it.

$$\forall x(\neg Gx \supset \neg Ex)$$

- *Not* the converse (any old way of growing it is explanatory).
- *Not* uniqueness (might be many *m's*).
- Generative sufficiency a necessary (but not sufficient) condition for explanation.
- ¬Furnish a Game in which the pattern is Nash
- ¬Furnish a Functional with respect to which the trajectory is an extremal
- ¬Furnish a Regression relating aggregate variables.

* … as against prediction.

Generative explanation*

- To explain a social regularity
- Demonstrate how it could emerge on time scales of interest to humans in a population of cognitively plausible agents
- Does the micro-specification *m* generate the macroscopic *explanandum x*
- If so, *m* is a generative explanatory candidate.
- Motto (Epstein, 1999) is negative : If you didn't grow it, you didn't explain it.

$$\forall x(\neg Gx \supset \neg Ex)$$

- *Not* the converse (any old way of growing it is explanatory).
- *Not* uniqueness (might be many *m's*).
- Generative sufficiency a necessary (but not sufficient) condition for explanation.
- ¬Furnish a Game in which the pattern is Nash
- ¬Furnish a Functional with respect to which the trajectory is an extremal
- ¬Furnish a Regression relating aggregate variables.

* … as against prediction.

Cognitively Plausible Agents

- Have emotions (notably fear)
- Have bounded deliberative capacity
- Have social connection
- And all of those might matter.
- Accordingly…

Agent_Zero

Endowed with distinct affective, deliberative, and social modules each grounded in contemporary neuroscience:

Internal modules interact to produce observable individual behavior.

Multiple agents interacting generate wide variety of collective dynamics: health, conflict, network dynamics, economics, social psychology, law.

Get synthesis started.

All provisional….

Old Idea…

- Hume: “Reason is a slave to the passions.”
- Aristotle/Spinoza: “Man is a social animal.”
- Looking for a simple convolution of:

$$\text{Passion} \oplus \text{Reason} \oplus \text{Social}$$

But Formal

Lots of empirical criticisms of the rational actor of Economics and Game Theory.

Gripes (even decisive experiments) do not change scientific practice.

Need explicit formal alternatives.

Albeit provisional, *Agent_Zero* is one: mathematical and computational.

- Before laying out the equations …

Big Picture…where we're going.
The violence interpretation

- Agents occupy an landscape of indigenous sites.
- There's a binary action agents can take: destroy some sites
- Experience produces a *disposition* to take the binary action
- Some sites are inactive/benign. Some active/fear-inducing
- Affect: Agents fear-condition on local stimuli: NOT decision-making of choosing.
    - Passion
- Bounded rationality: Local sample relative frequency
    - Reason
- Add these. Solo Disposition (propensity to perform the act).
- Social animals: Add others' weighted Solo Dispositions
    - Weights are endogenous (minimize parameters)
- If Total Disposition exceeds threshold, take the action.
    - Destroy, or flee, or refuse vaccine, or dump assets, or find guilty…

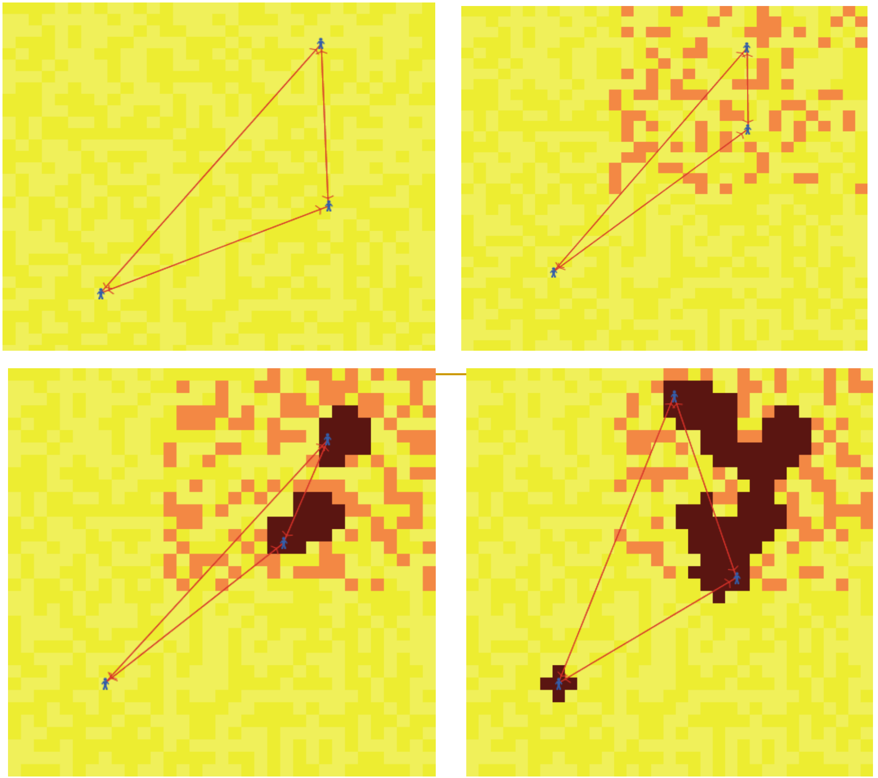
Computational **Parables** : Slaughter of Innocents

Vision Von Neumann

Agent #0 fixed in SW: zero direct stimulus

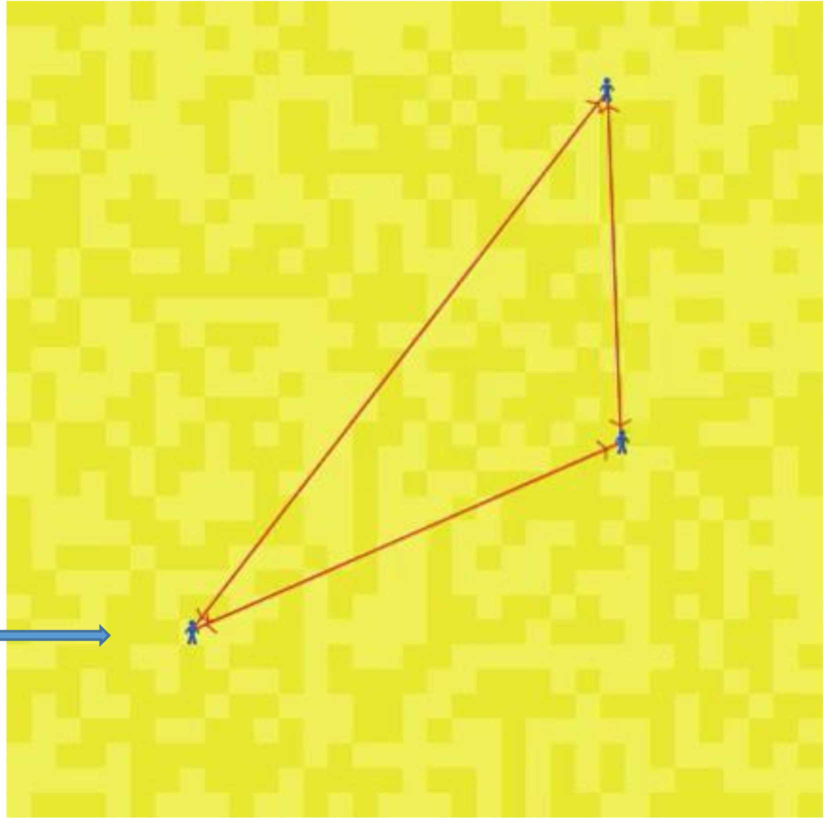
Others in NE: stimulus, violent action

By dispositional contagion, Agent 0 acts.



Parable 1: Agent_Zero Joins Without Direct Stimulus (eye candy runs are just sample paths, of course)

Since no stimulus within sensory radius. Would not act alone



Overall Set-Up

Action , Threshold

- Binary Action $A \in \{0,1\}$
  - Flee snake or don't
  - Raid icebox or don't
  - Join lynching or don't
  - Refuse vaccine or don't
  - Dump stock or don't
  - Wipe out village or don't
  - "Behavior" will mean a binary action.
- Nonnegative Real Threshold $\tau \geq 0$

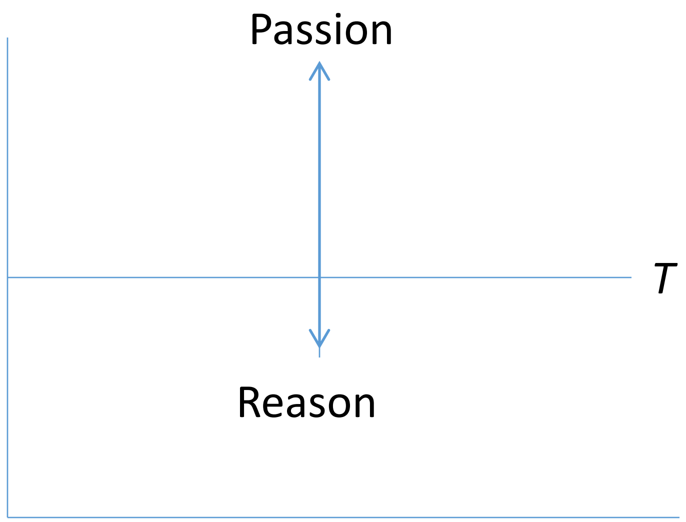
Solo Disposition to Act

- Agents endowed with Affective V(t) and Deliberative real-valued functions P(t) bounded to [0,1] defined on a stochastic stimulus space, and each Solo disposition is, for the moment, as simple as possible, their sum:

$$D_i^{solo}(t) = V_i(t) + P_i(t)$$

- Addition also nice when they compete given a threshold

The “Hume Equation”



$$(P - T > T - R) \supset (P + R > 2T)$$

Or, if you prefer

Average of Passion and Reason Exceeds Threshold

But Connected: Total Disposition to Act

- Again, solo disposition is the sum: $D_i^{solo}(t) = V_i(t) + P_i(t)$
- But Agents also carry weights (unconsciously I presume): $\omega_{ji}(t)$
- We therefore define the Total Disposition to Act as*

$$D_i^{tot}(t) = D_i^{solo}(t) + \sum_{j \neq i} \omega_{ji}(t) D_j^{solo}(t)$$

*self-weights assumed to be one, but can relax (low self-esteem agents).

Action Rule

Act if and only if $D_i^{tot}(t) > \tau_i$.

Or, defining $D_i^{net}(t) \equiv D_i^{tot}(t) - \tau_i$,

The Action Rule reduces to:

Act iff $D_i^{net}(t) > 0$.

Dispositional Contagion, Not Imitation of Behavior

- Nobody else's observable *action* appears in this equation.
- Hence, the mechanism of action cannot be imitation of behavior, because the binary acts of others are not registered in this calculation.
- So we are suspending a "monkey-see/monkey-do" assumption central to much literature on social transmission.
- Obvious problem with imitation of observable action: no mechanism for the *first* actor. Nobody to imitate.
- (Noise is cheating…not a mechanism)

Under The Hood: Provisional Parsimonious

Act if $D_i^{tot}(t) = V_i(t) + P_i(t) + \sum_{j \neq i} \omega_{ji}(t)(V_i(t) + P_i(t)) > \tau_i$

τ_{i}'s equal, so one parameter

$v_i(t)$ solves $\frac{dv_i}{dt} = \alpha_i \beta_i v_i^{\delta}(\lambda - v_i)$ Nonlinear Rescorla - Wagner

Original is just $k_i(1 - v_i)$. So also one parameter. Neuroscience coming up.

$P_i(t, x; m) = \frac{1}{m} \sum_{t-m}^{t} RF(x)$ Moving average of local relative frequency

Memory a third parameter. Weights endogenous

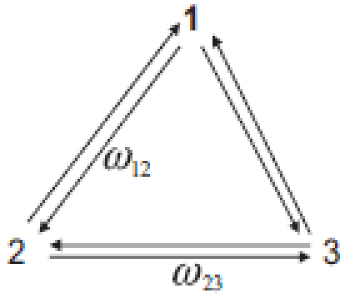
$\omega_{ji}(t) = [v_i(t) + v_j(t)](1 - \left|v_i(t) - v_j(t)\right|)$ Strength - scaled affective homophily

Where t meters trials for mobile agents on a spatial stimulus landscape

Humble goal:

- Get the synthesis started
- Provisional plausible/testable modules

ODE and ABM versions
(Math and *Mathemtica* Code in the book)



$$\frac{dv_1}{dt} = \alpha_1 \beta_1 v_1^{\delta_1} (\lambda - v_1) \qquad [25]$$

$$\frac{dv_2}{dt} = \alpha_2 \beta_2 v_2^{\delta_2} (\lambda - v_2)$$

$$\frac{dv_3}{dt} = \alpha_3 \beta_3 v_3^{\delta_3} (\lambda - v_3)$$

$$v_i(0) = v_0$$

Weights define dispositional network. Extract v-functions and compute net dispositions:

$$D_1^{\text{net}}(t) = v_1(t) + P_1 + \omega_{21}(v_2(t) + P_2) + \omega_{31}(v_3(t) + P_3) - \tau_1 \qquad [26]$$

$$D_2^{\text{net}}(t) = v_2(t) + P_2 + \omega_{12}(v_1(t) + P_1) + \omega_{32}(v_3(t) + P_3) - \tau_2$$

$$D_3^{\text{net}}(t) = v_3(t) + P_3 + \omega_{13}(v_1(t) + P_1) + \omega_{23}(v_2(t) + P_2) - \tau_3$$

FIGURE 24. Generalized Three-Agent Model

The Subtitle of *Agent_Zero*

- Toward Neurocognitive Foundations for Generative Social Science
- Talked about Generative Social Science
- What’s this neurocognitive business?

Fear Instantiation

- Thinking of cycles, we care about:
- Fear acquisition
- Fear extinction

- Now I will butcher some neuroscience…

Amygdala Circuit

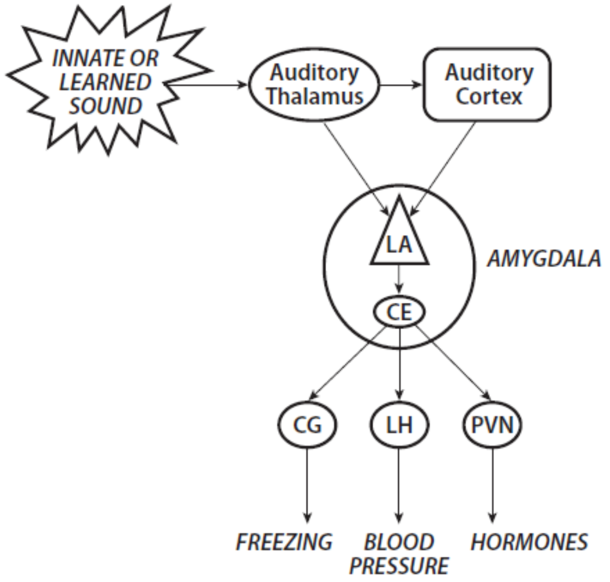


FIGURE 1. Auditory Amygdala Stimuli and Defense Responses. Source: LeDoux (2002, Figure 5.6)

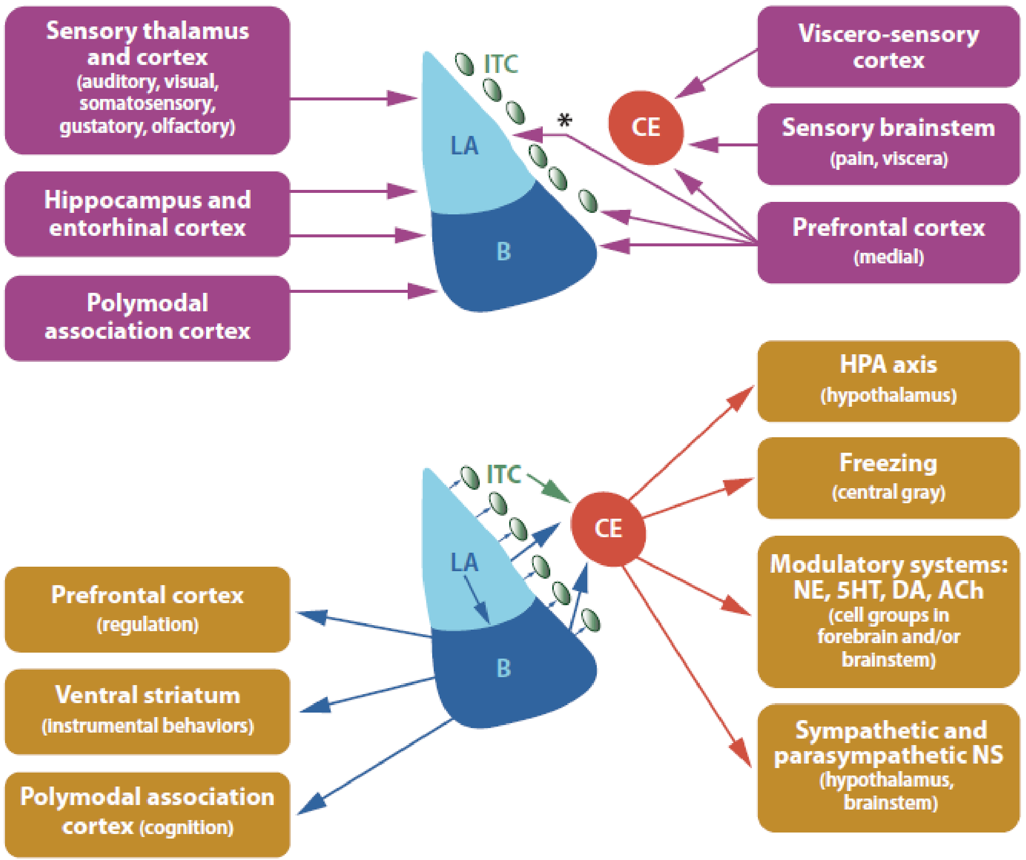


FIGURE 2. Amygdala Inputs and Outputs. Inputs to some specific amygdala nuclei. Asterisk (*) denotes species difference in connectivity. (Bottom) Outputs of some specific amygdala nuclei. 5HT, serotonin; Ach, acetylcholine; B, basal nucleus; CE, central nucleus; DA, dopamine; ITC, intercalated cells; LA, lateral nucleus; NE, norepinephrine; NS, nervous system. Source: Rodrigues, LeDoux, and Sapolsky (2009)

Amygdala Areas: Various Stains

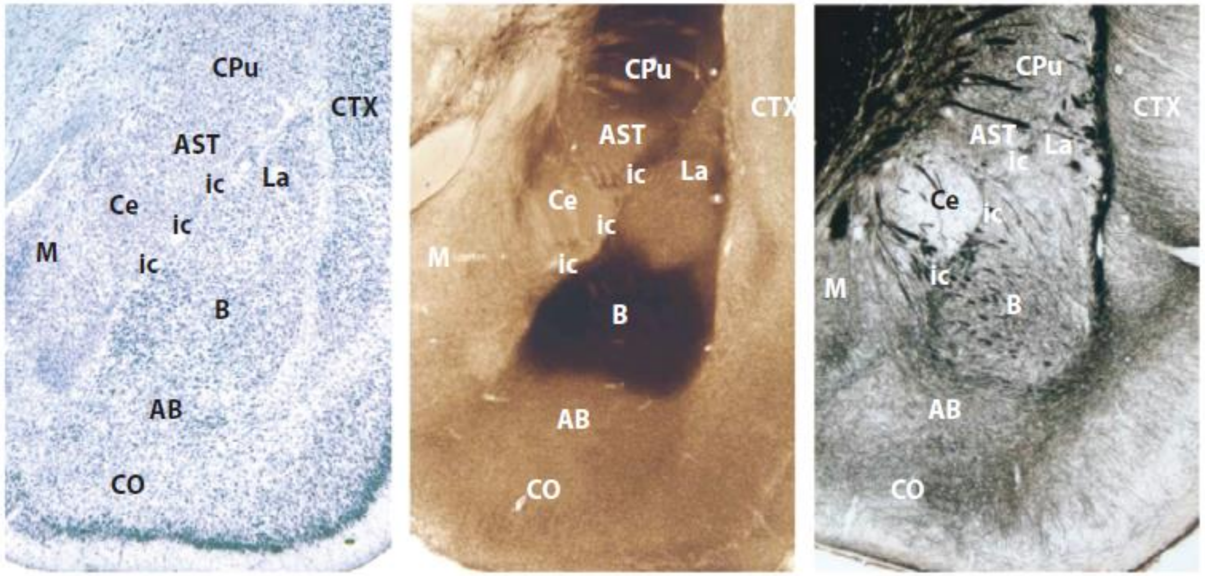


FIGURE 3. Key Areas of the Amygdala. Key areas of the amygdala, as shown in the rat brain. The same nuclei are present in primates, including humans. Different staining methods show amygdala nuclei from different perspectives. Left panel: Nissl cell body stain. Middle panel: acetylcholinesterase stain. Right panel: silver fiber stain. Abbreviations of amygdala areas: AB, accessory basal; B, basal nucleus; Ce, central nucleus; itc, intercalated cells; La, lateral nucleus; M, medial nucleus; CO, cortical nucleus. Non-amygdala areas: AST, amygdalo-striatal transition area; CPu, caudate putamen; CTX, cortex. Source: LeDoux (2008, p. 2698); reprinted courtesy of Joseph E. LeDoux

Don’t Care *Where*…Care that it’s Innate, Automatic, Fast, Inaccessible to Deliberation

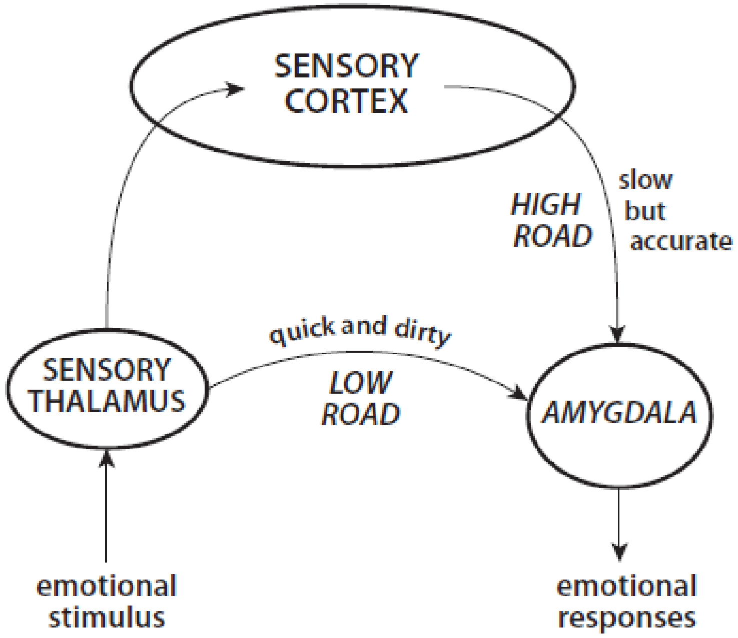


FIGURE 4. Low Road and High Road to Fear. Source: LeDoux (2002, pp. 61–63, Figure 5.7)

Also equipped with an associative machinery.

“Neurons that fire together wire together.” Donald Hebb (1949)

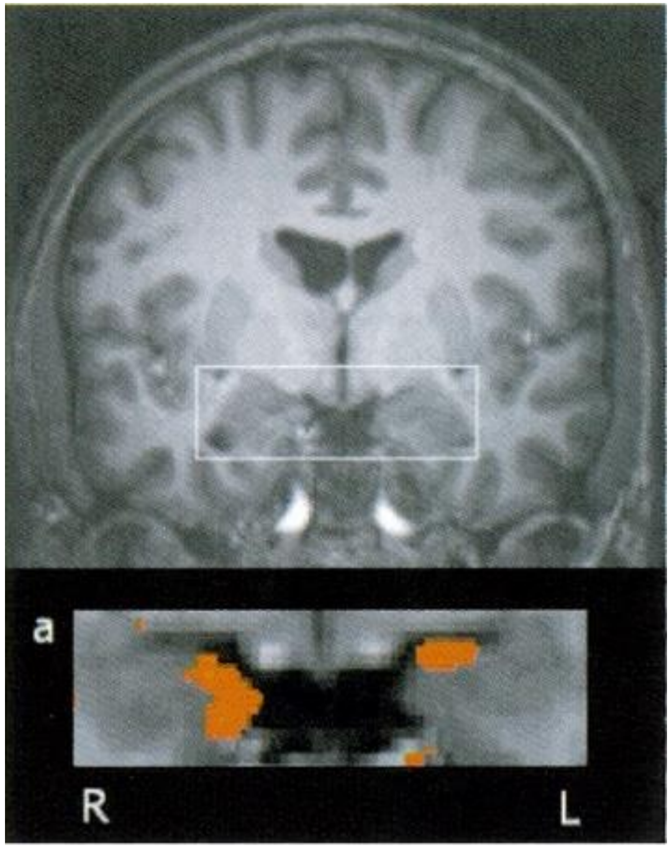
Associative Fear Conditioning: Acquisition Phase

US: Shock cuff

UR: Amygdala activation

CS: Blue Light (neutral)

CS-US Pairing Trials
    Light…Shock
    Light…Shock
    Light…Shock
    Light alone …………→



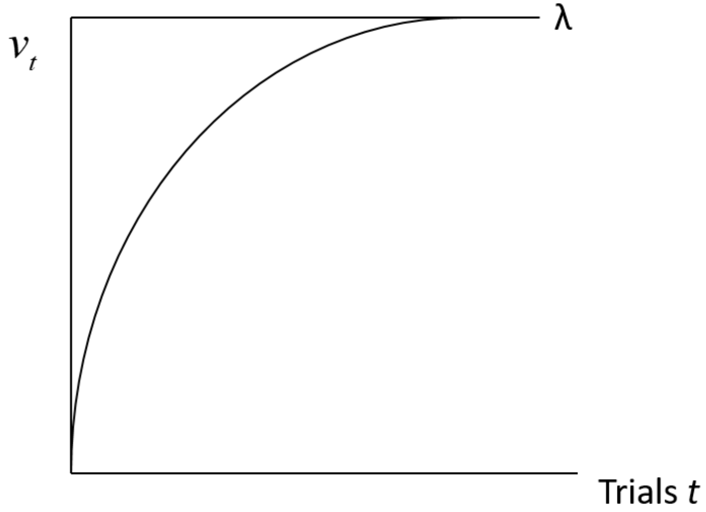
Simple Elegant Model of Associative Learning
Rescorla-Wagner Model (1972)

$$v_{t+1} - v_t = \alpha\beta(\lambda - v_t)$$

Learning rates (α, β) : Surprise and Salience

Associative gain requires Surprise and Salience

λ (typically 1) is max associative strength.



Important

- NOT modeling brain regions or tissue.
- Modeling an innate associative performance conferred by the neural architecture and
- Explained by the underlying neuroscience.
- The neuroscience ‘licenses’ the modeling and its interpretation.
- …can now explain what Hume observed.

Hume’s Association of Ideas

“. . . after the constant conjunction of two objects . . . we are determined by *custom* alone to expect the one from the appearance of the other . . . Having found in many instances, that two kinds of objects—flame and heat, snow and cold—have always been conjoined together; if flame or snow be presented anew to the senses, the mind is carried by *custom* to expect heat or cold.**”** It is not by reasoning, moreover, that we form the connection. “All these operations are a species of natural instinct, which no reasoning or process of the thought and understanding is able either to produce or to prevent” (Section V, Part I).

Very important: Nonconscious and inaccessible to ratiocination.

Perils of Fitness

- "Survival circuits" (LeDoux 2012) conserved across vertebrate evolution.
- Epstein (2013) "Pleistocene man never encountered a BMW, but *we* freeze when a car whips around the corner at us, just as *he* froze when huge animals charged suddenly from the tall brush. *We are harnessing the same innate fear-acquisition capacity—the same innate neurochemical computing architecture.* Miraculously, synaptic plasticity permits us to adapt the evolved machinery to encode novel threats."
- Invaluable but very dangerous…double-edged

Surprise + Salience → Strong Conditioning

CS	US	UR/CR
Light	Shock	Fear
Vietnamese Face	Ambush	My Lai
Arab Face	9/11	Koran as ISIS Field Manual
Japanese Face	Pearl Harbor	Internment

Surprise + Salience → Strong Conditioning

CS	US	UR/CR
Light	Shock	Fear
Doctor	Tuskegee	Distrust
MMR Vaccine	Autism	Vaccine refusal
Financial asset	Sudden devaluation	Panic

Also Over-General and Persistent

- *Should* stay afraid of hippos.
- Affect can remain above the threshold long after actual stimulus has stopped.
- Stimulus stopped at t. Extinction may be far off. Extreme case is PTSD.
- Not Symmetrical

Full Affective Trajectory (vigilance and complacency)

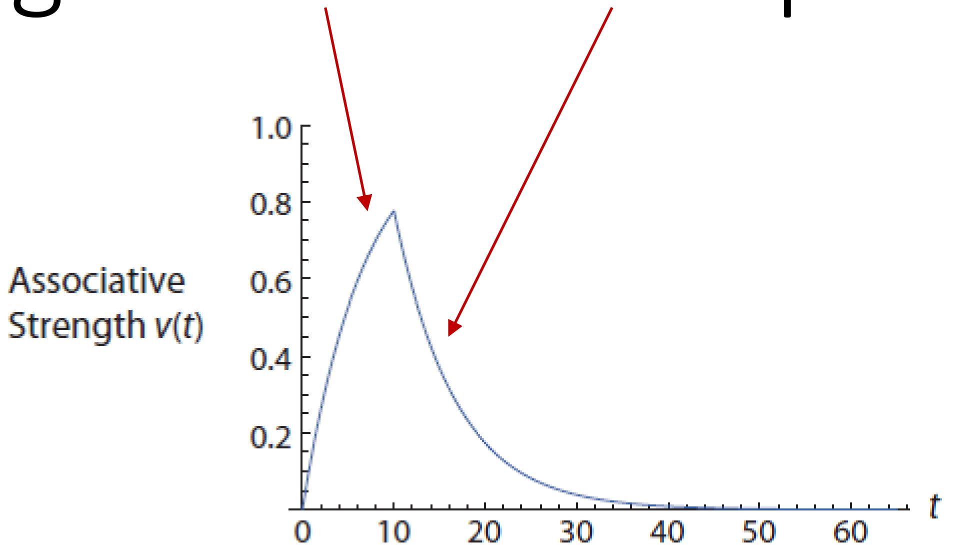
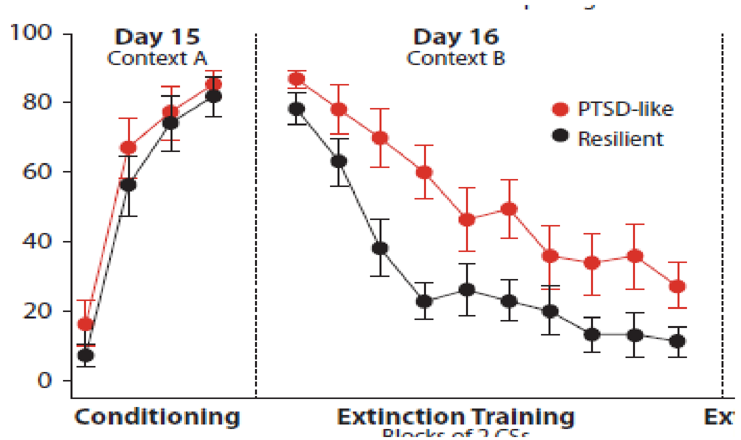


FIGURE 8. Acquisition and Extinction



Rats, predatory threat

With t^* the time at which trials cease, the full solution is then

$$v(t) = \lambda(H(t^* - t)(1 - e^{-\alpha\beta t}) + (1 - e^{-\alpha\beta t^*})H(t - t^*)e^{-\alpha\beta(t - t^*)})$$

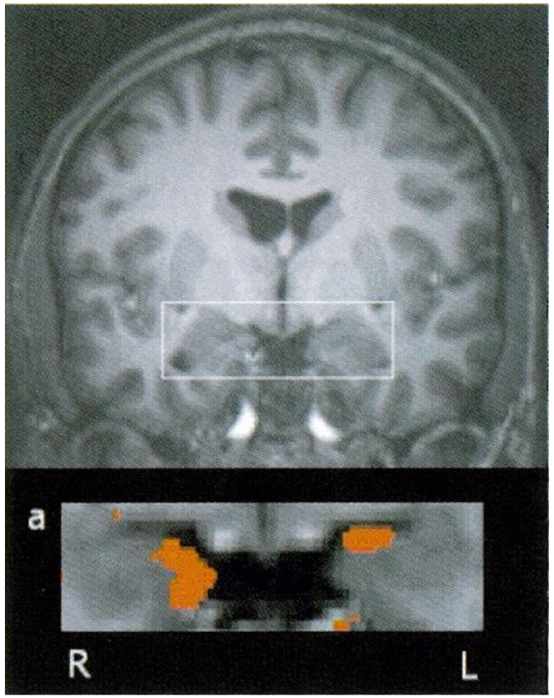
We do not fear *what* the rat fears, but we fear *how* the rat fears.

One Component is Fear, But Saying Contagious (ω)

- Fear contagion discussed by Adam Smith, Gustav LeBon…Financial "panics"
- Coupled Contagion model shows it's a fruitful modeling postulate.
- But any neural basis/license?
- Yes!

Observational Fear Conditioning*

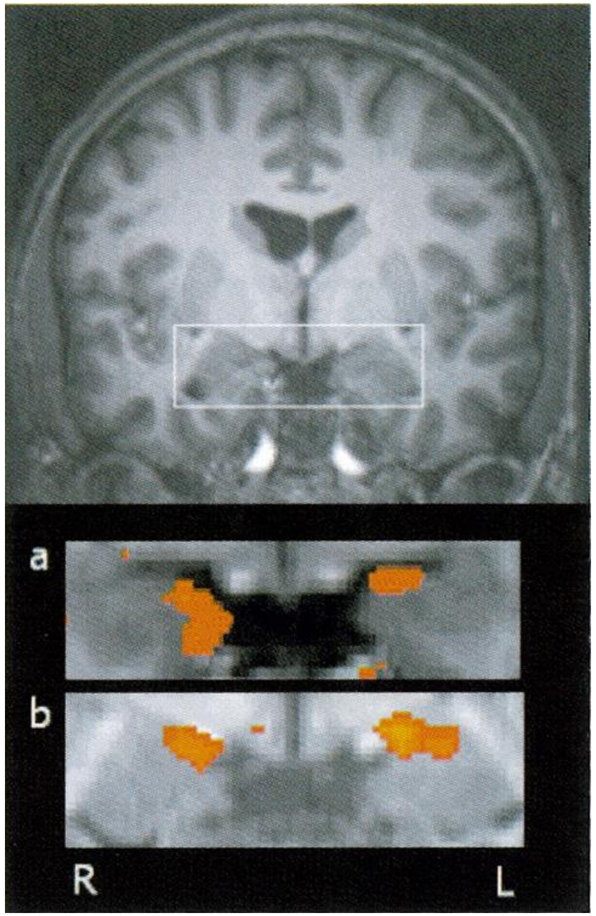
- Shown earlier : Fear-Conditioned human amygdala fMRI
- US: Shock cuff
- UR: Amygdala activation
- CS: Blue Light (neutral)
- CS-US PairingTrials
  - Light…Shock
  - Light…Shock
  - Light…Shock
  - Light alone ………….→



*Olsson, A., Nearing, K. I., & Phelps, E. A. (2007). Learning fears by observing others: the neural systems of social fear transmission. *Social cognitive and affective neuroscience*, *2*(1), 3–11.

Is Fear Contagious?

- Top Panel (a), fMRI of subject above
- True Subject: Bottom Panel (b), fMRI of *observer.*
- Watches the blue-shock pairings
  - Then is shown blue light alone…
  - Same fMRI as if conditioned!
- Advantage clear
  - I learn to fear the fire by watching you get burned
  - Downside is also clear: rapid nonconscious transmission of baseless fear.



Ingredient 1: Emotion

- Introduce a generalized version of the classic (1972) Rescorla-Wagner model and emotional contagion through weights (endogenous functions of affect in book).

Reason may be "a slave to the passions," but once in a great while, it happens…however badly!

- Typically we have incomplete and imperfect information
- Make systematically erroneous appraisals of it.
- Robustly documented errors:
  - Framing effects (medical decisions)
  - Endowment effects (loss aversion)
  - Representativeness heuristic
    - Local sample represents population
  - Base rate neglect
    - Confuse $P(+|sick)$ with $P(sick|+)$
  - Anchoring on what you hear first
    - $2345678 < 8765432$
- *Agent_Zero* (local relative frequentist) exhibits several.

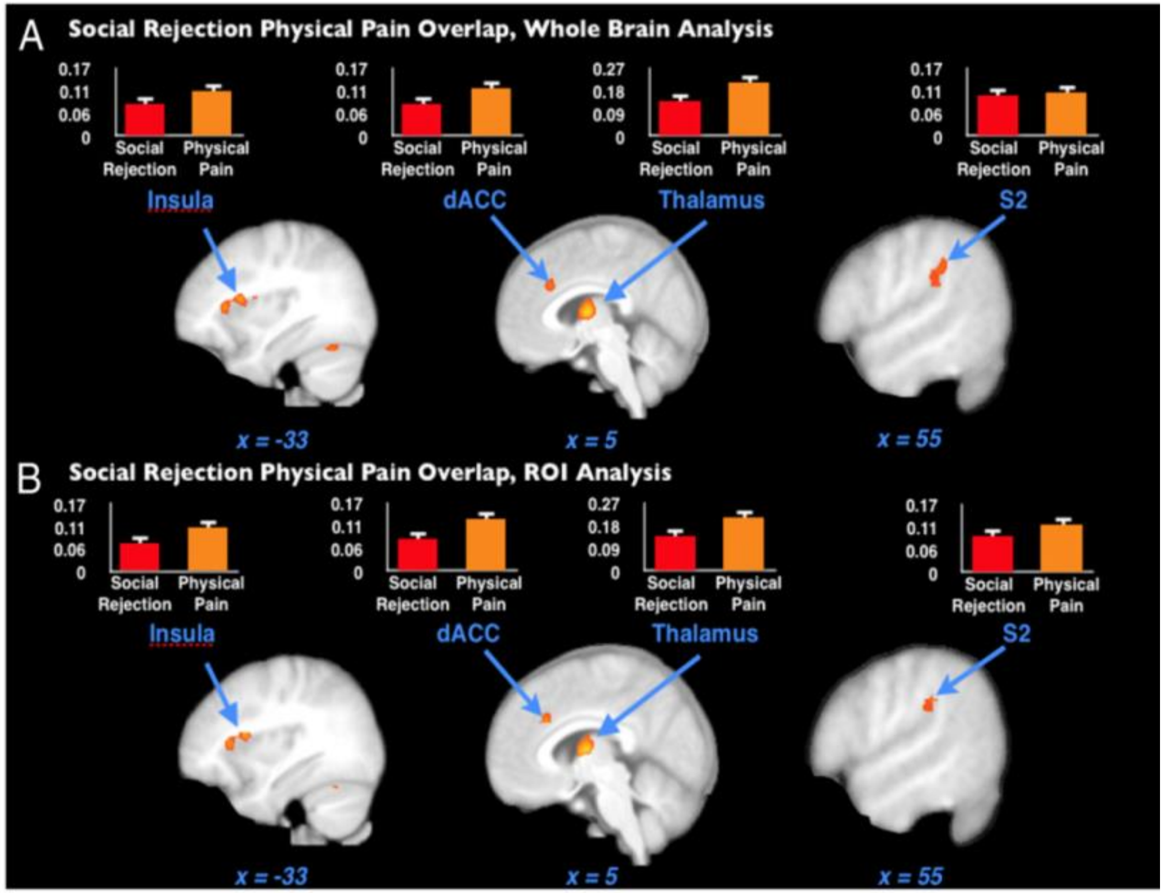
To Make Matters Worse…

- Agents driven by strong (unconscious) emotions (like fear), doing bad statistics on incomplete and biased data, also *influence* one another.
- Conformist pressures can then produce widespread convergence on counter-productive behavior.
- Conformity effects are documented in many spheres (since Asch 1958).
- Again, a neural basis?

Yes: Nonconformity Hurts!

- Kross et al (*PNAS* 2011) "…when rejection is powerfully elicited…areas that support the sensory components of *physical pain* (secondary somatosensory cortex; dorsal posterior insula) become active."
- Illustrated in fMRIs below.

Neural Drivers of Conformity



Neural overlap between social rejection and physical pain.

Bar graph: no statistically significant difference between (βs of) rejection and physical pain. **Positive predictive value = 88%.**

Source: Kross, E., Berman, M. G., Mischel, W., Smith, E. E., & Wager, T. D. (2011). Social rejection shares somatosensory representations with physical pain. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(15), 6270–6275.

Conform Because Rejection Hurts.

- As they write, “These results give new meaning to the idea that rejection ‘hurts’…rejection and physical pain are similar not only in that they are distressing—they share a common somatosensory representation as well.”
- We give others weight…so

Ingredient 3: Network Weights

- Agents experience a weighted sum of the affective and deliberative states of others
- As discussed, weights are actually endogenous in model—strength-scaled affective homophily *generates* networks…more on this below.

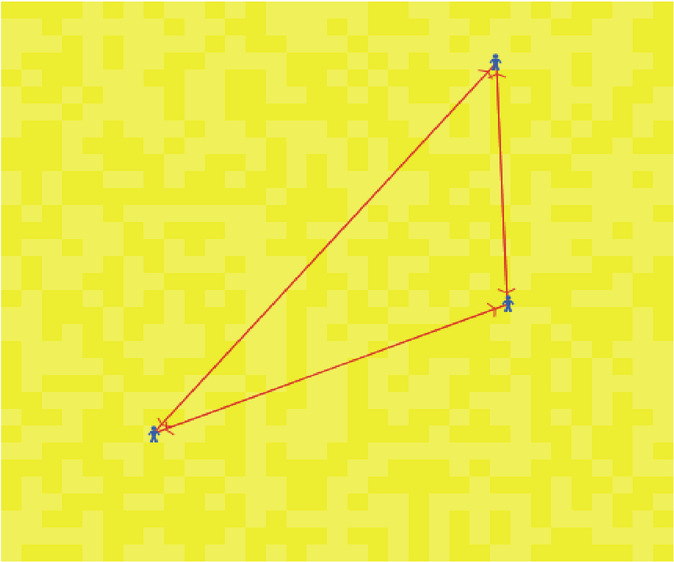
Given these components…

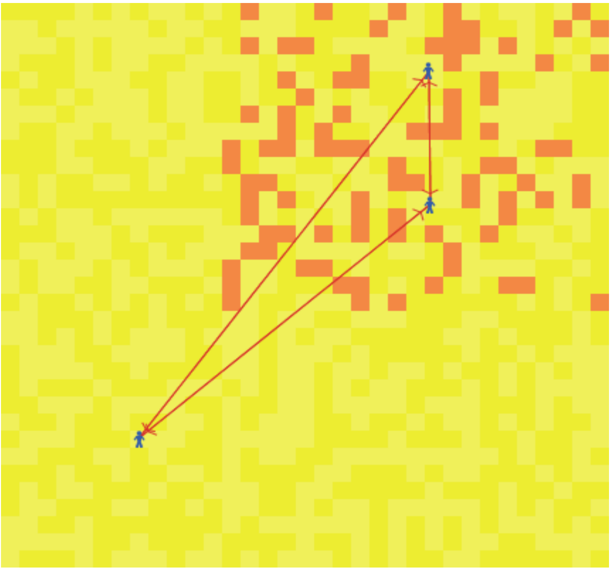
- Logic of the Model:
  - Disposition
  - Threshold
  - Action
    - This typically alters the stimulus pattern

Agent-Based Model Runs : Computational Parables

All Code for ODEs and for the ABM, all parameter values, all initial conditions and random seeds are at the Princeton Press Agent_Zero site. Replicable.

Landscape and Trials: Agent_0 Fixed and Mobile Rovers

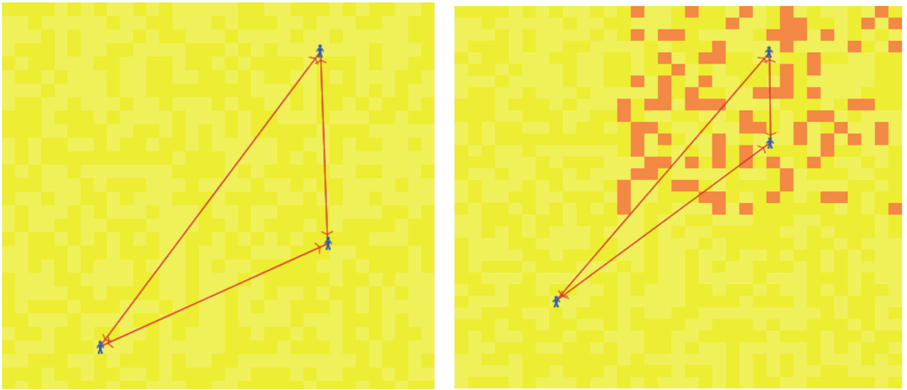




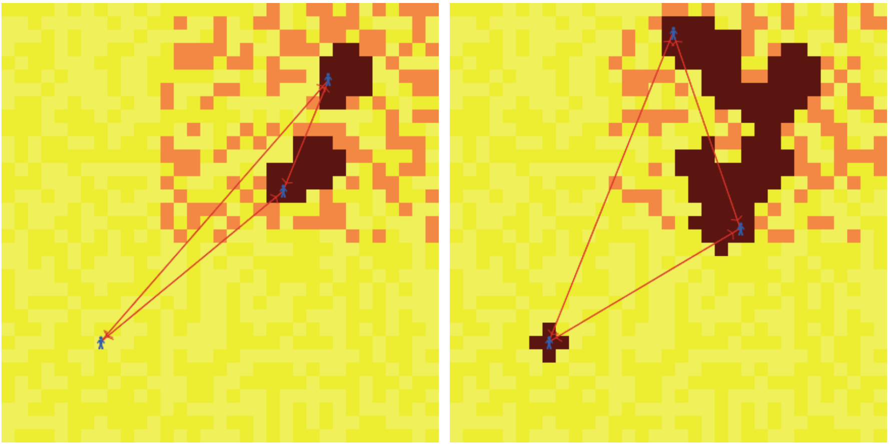
Agents directly condition on orange trials and compute RF w/in vision. Then a weighted sum over network. If D>τ, destroy all sites w/in damage radius

Parable 1: Slaughter of Innocents

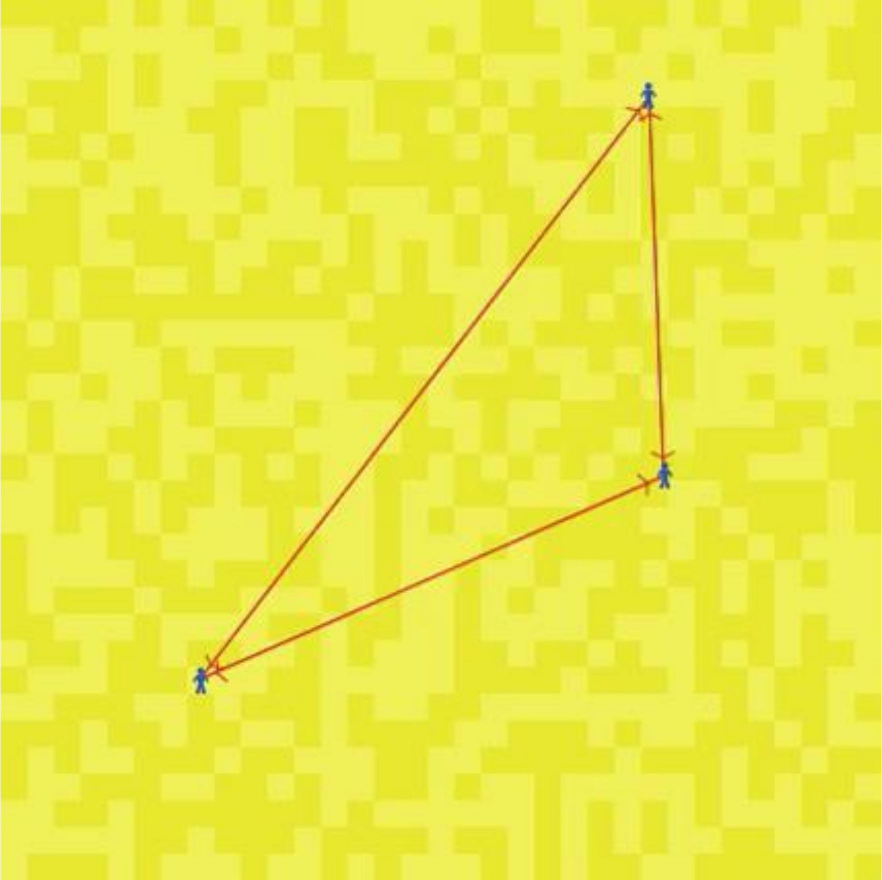
Agent 0 fixed, zero direct stimulus
Mobile rovers transmit retaliatory disposition
Vision Von Neumann…..Agent 0 massacres village



Animation 3.3. Activation by Dispositional Network Effect:



Parable 1: Agent_Zero Joins Without Direct Stimulus



V=P=0, since no stimulus within sensory radius

Solo disposition = 0

Eye candy is one sample path. Turn off and build statistical portrait.

La Condition Humaine

- Why?

$$D_i^{tot}(t) > \tau > D_i^{solo}(t)$$

- You take action in group (since $D_i^{tot}(t) > \tau$) that you would not take alone (since $\tau > D_i^{solo}(t)$).

- Indeed, you may be the only agent with this ordering. In that case:

- *Despite being negatively disposed* you act first!*

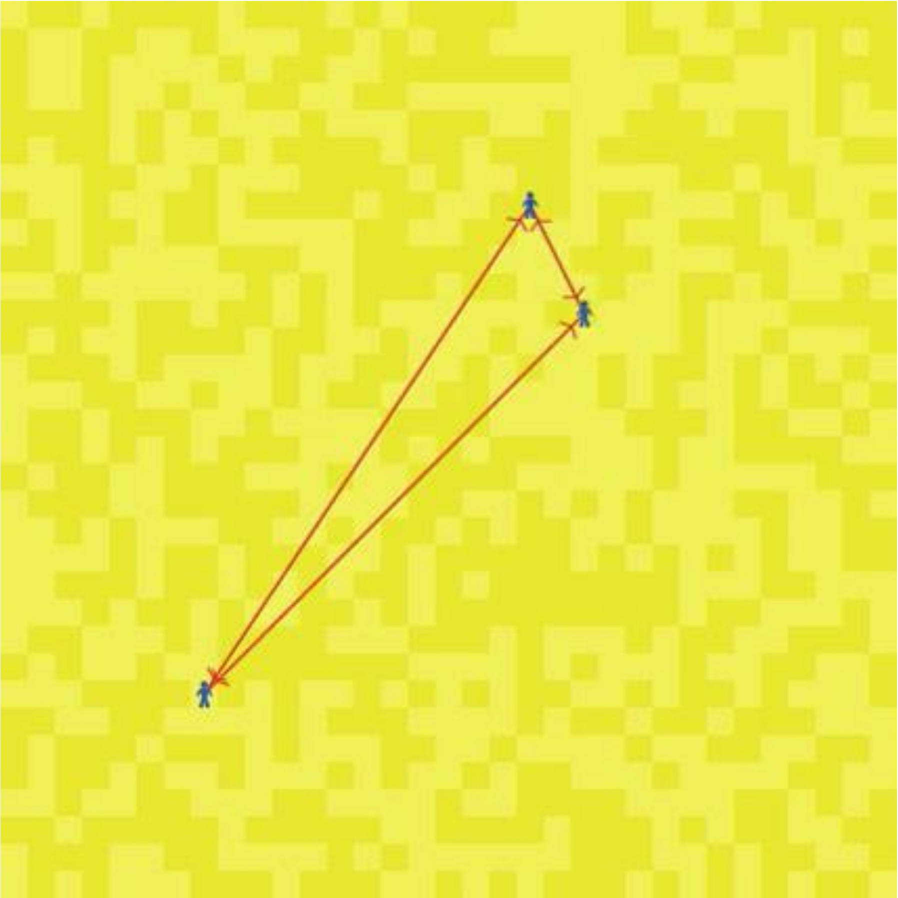
* $D_i^{solo}(t) - \tau < 0$

Parable 2: Agent Zero Initiates

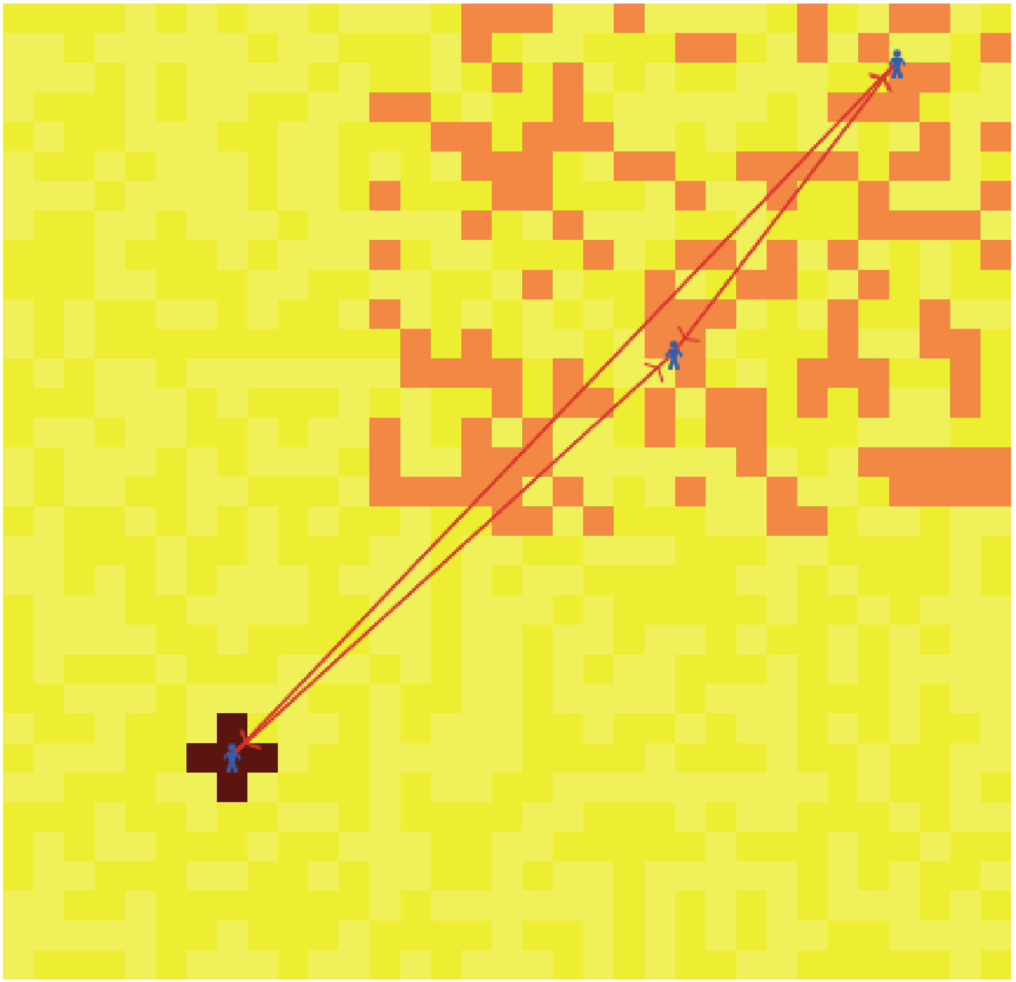
- Again, ***no*** direct stimulus

- He goes first!

- Not imitation of behavior



Core Parable: Agent_Zero Goes First Without Stimulus



Leadership or Susceptibility?

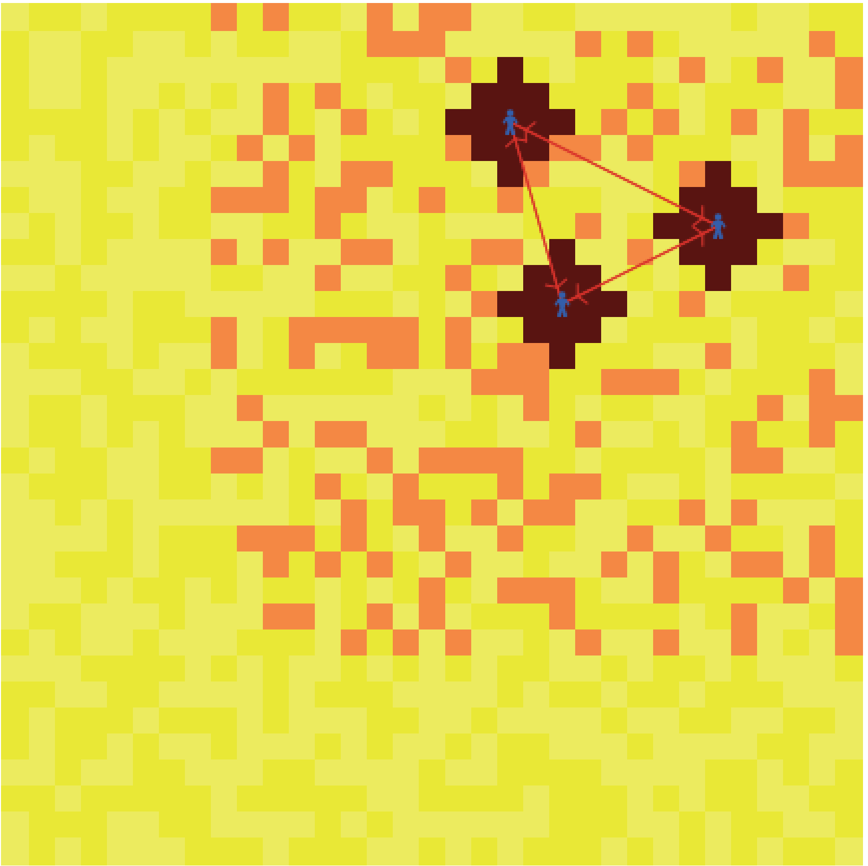
- Not behavioral imitation.
  - If 1st, nobody to imitate
- Leader, or just most susceptible to D-contagion?
- Tolstoy's answer: 'A king is history's slave, performing for the swarm life.' (*War and Peace*, 1896)

Unsettling Picture

"The overall picture of Homo sapiens reflected in these interpretations of *Agent_Zero* is unsettling: Here we have a creature evolved (that is, selected) for high susceptibility to unconscious fear conditioning. Fear (conscious or otherwise) can be acquired rapidly through direct exposure or indirectly, through fearful others. This primal emotion is moderated by a more recently evolved deliberative module, which, at best, operates suboptimally on incomplete data, and whose risk appraisals are normally biased further by affect itself. Both affective and cognitive modules, moreover, are powerfully influenced by the dispositions of similar—equally limited and unconsciously driven—agents. Is it any wonder that collectivities of interacting agents of this type—the *Agent_Zero* type—can exhibit mass violence, dysfunctional health behaviors, and financial panic?" (Epstein, 2013)

Fight vs. Flight

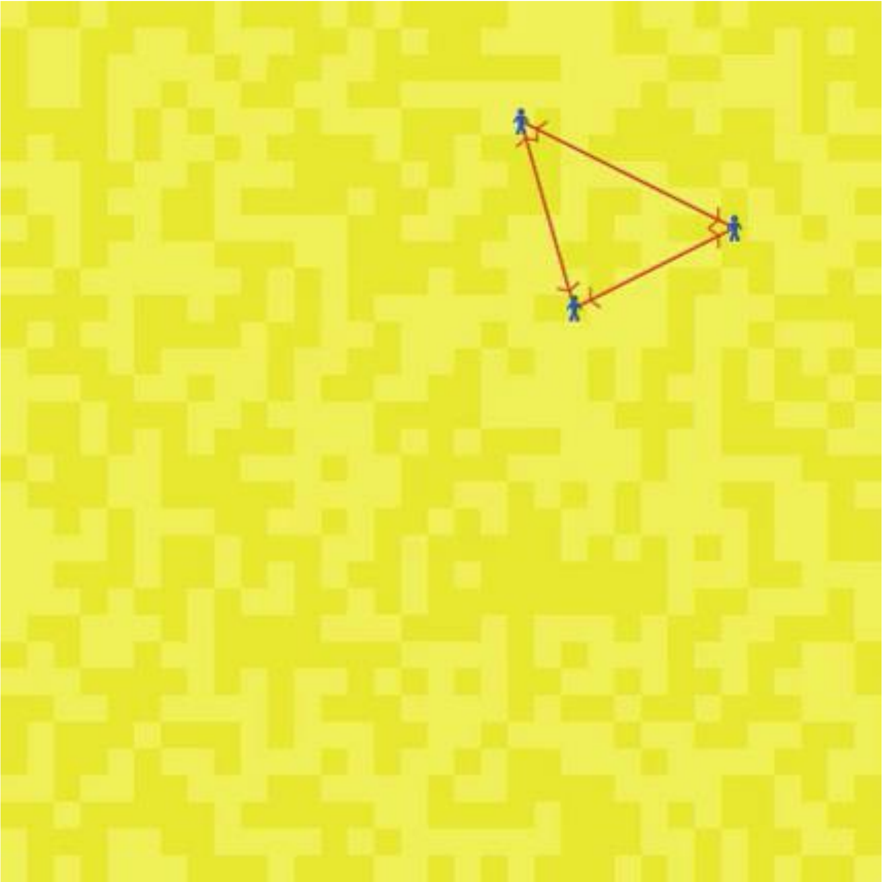
- Fight



Flight

- Katrina Evacuees
- Syrian Refugees
- Capital/Portfolio Flight

- Recalcitrant agents “dragged out” by others.



The Reverse: Bystander Effects

- The classis experiment here is : Darley, J. M., & Latane, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology, 8*(4, Pt.1), 377-383.
- Smoke under the door experiment (1968)
  - Subject alone in the room leaves quickly.
  - Subject with 2 oblivious confederates in the room takes 2-3 times as long.
  - *Agent_Zero* exhibits this…
- Erez Hatna and I are analyzing these parables systematically and under many variations to determine robustness.
- But these “existence” results are interesting.

In all of this, Networks are Implicated

- How do network weights change?

- Why do networks happen?

Endogenous Weight Change by Affective Homophily (so weights are not parameters)

- Affective homophily. Affects changing. So try: $\left|v_i(t) - v_j(t)\right|$
- Problem: equals zero when identical; want 1.0 when equal.
- OK, so as homophily, use: $1 - \left|v_i(t) - v_j(t)\right|$
- Problematic as a weight: nudniks (v=0) same strength as crusaders (v=1). So, scale by total strength

$$\omega_{ji}(t) = [v_i(t) + v_j(t)](1 - |v_i(t) - v_j(t)|)$$

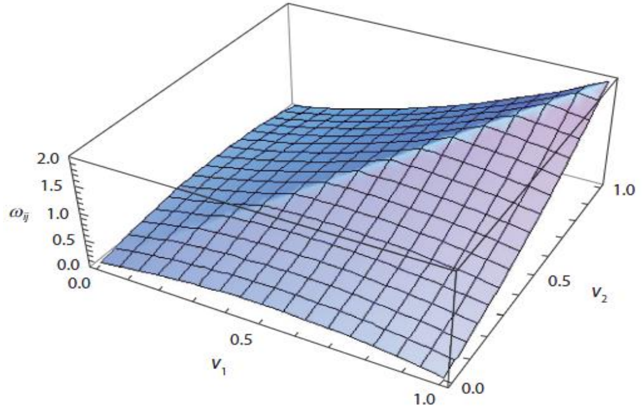
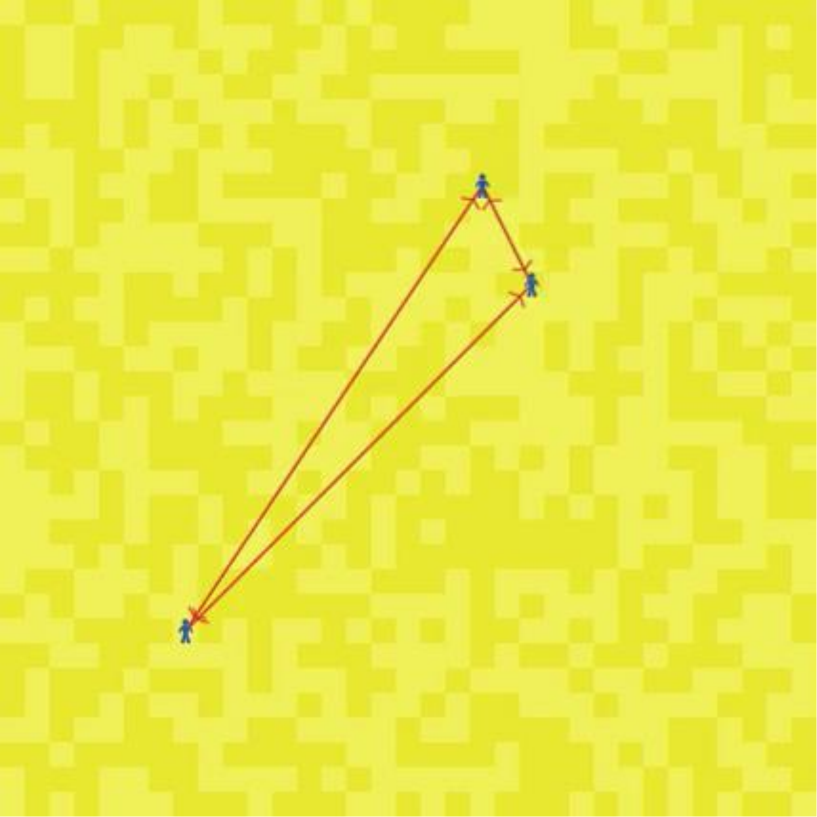


FIGURE 63. Weight Surface

Lazer, David. "The co-evolution of individual and network." *Journal of Mathematical Sociology* 25.1 (2001): 69-108.
Lazer, David, et al. "The coevolution of networks and political attitudes." *Political Communication* 27.3 (2010): 248-274.

Endogenous Weights:



- Affective homophily strengthens connection
- And this can matter immensely…

Grow The Arab Spring
Case 1: No Communication

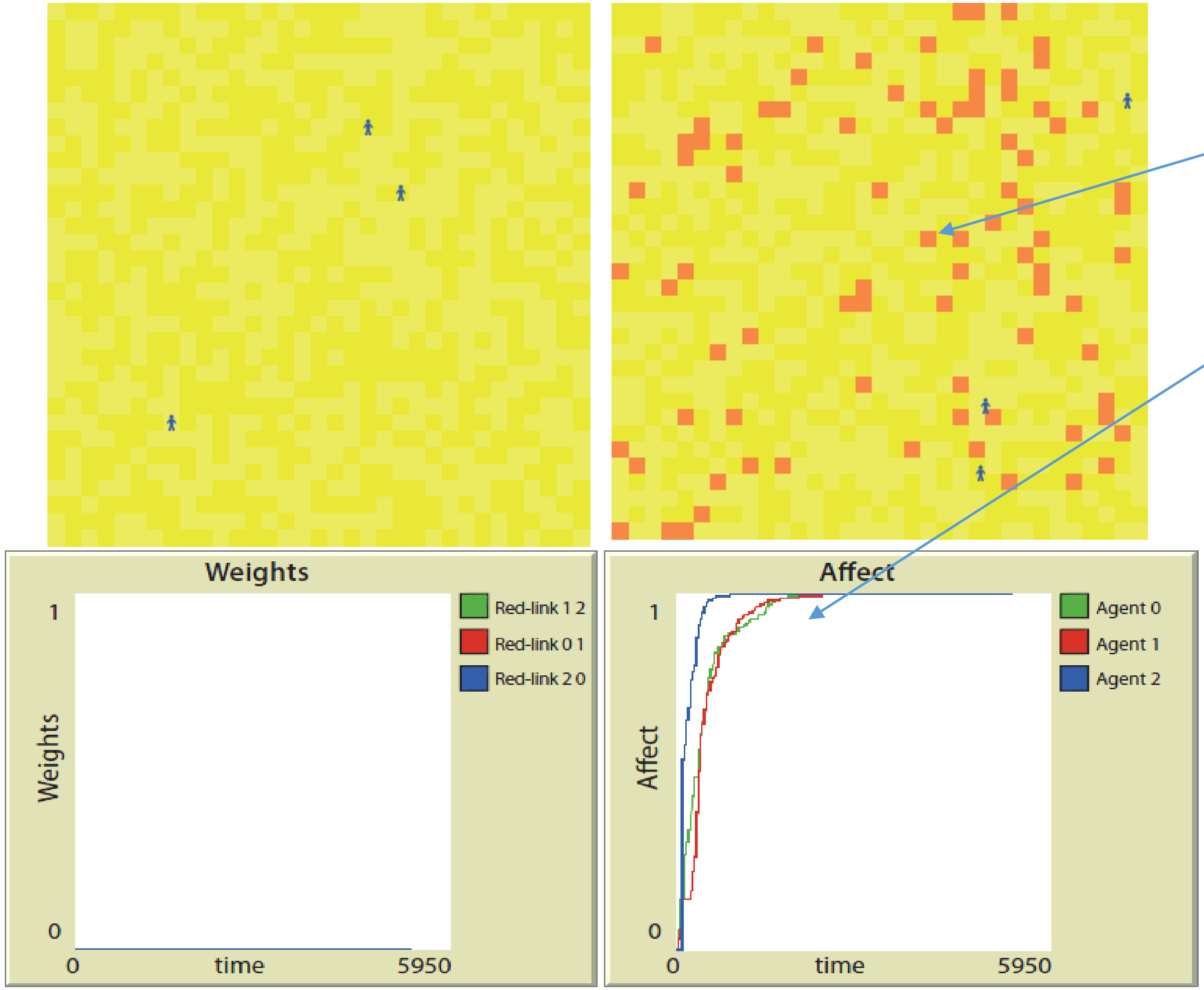


FIGURE 73. Weights and Affect with No Social Media

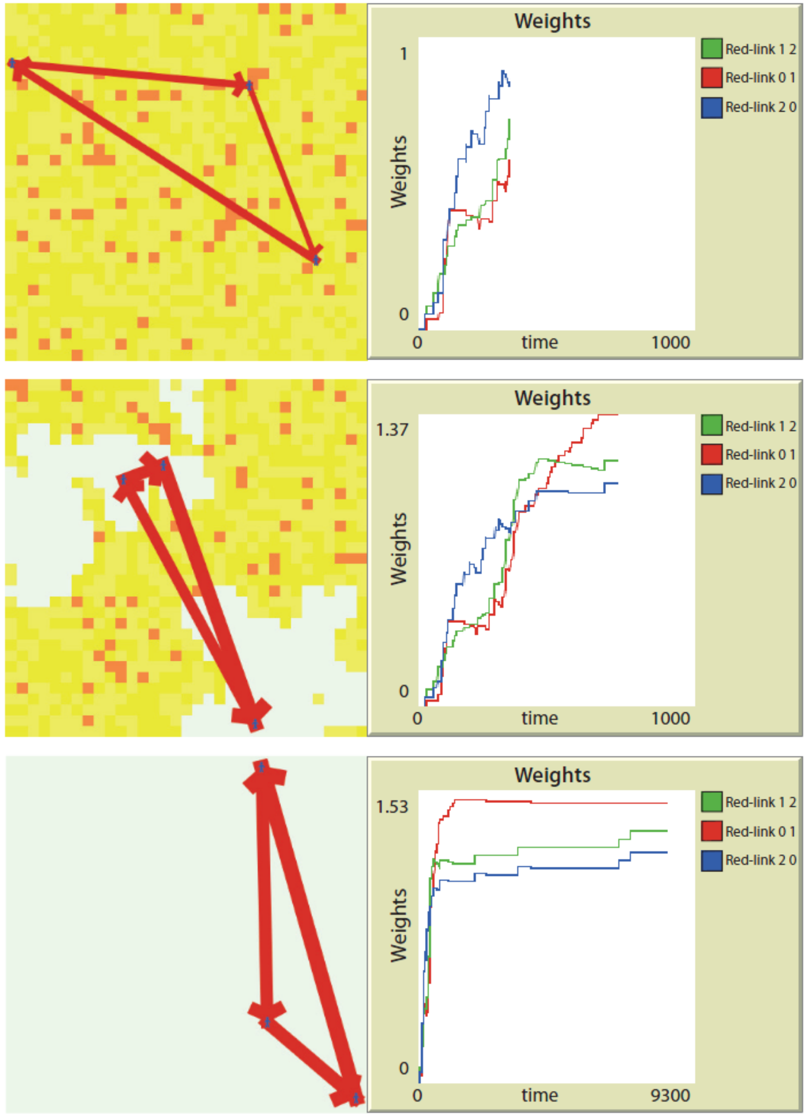
Instances of regime corruption (abduction, torture, theft, civil liberties)

Produce profound grievance

Weights clamped at zero by Big Brother.

In isolation, no action.

Arab Spring (Jasmine Revolutions)
Case 2: Communication→Dispositional Amplification→Overthrow



Revolt of the Swarm

- Leaderless Revolutions
  - No Mao, Lenin

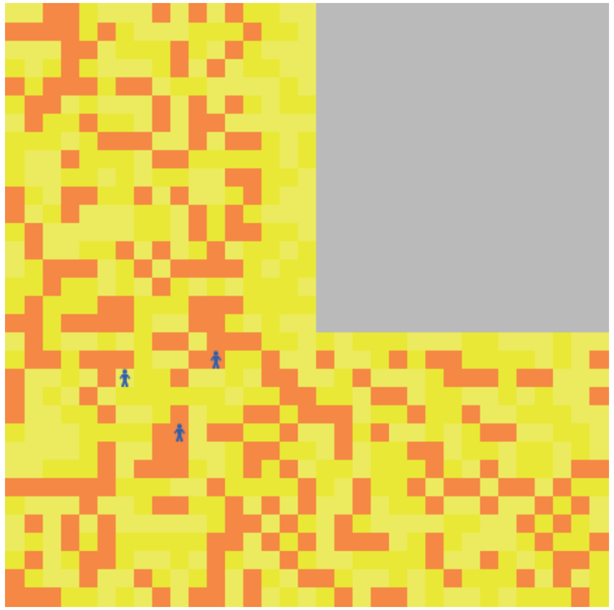
- Similarly in Juries

Jury Dynamics: 12 Angry Agent_Zeros

- Pre-Trial: General landscape of stimuli about OJ's guilt. Initial dispositions to convict are formed. Jurors strangers. All weights off.
- Trial: Competing stimuli (Prosecution and Defense). Dispositions are updated. Jurors do not communicate. Weights still off.
- Sequestration: Now homophily dynamics and network effects operate strongly.
- Agents convict in group when they would acquit alone:

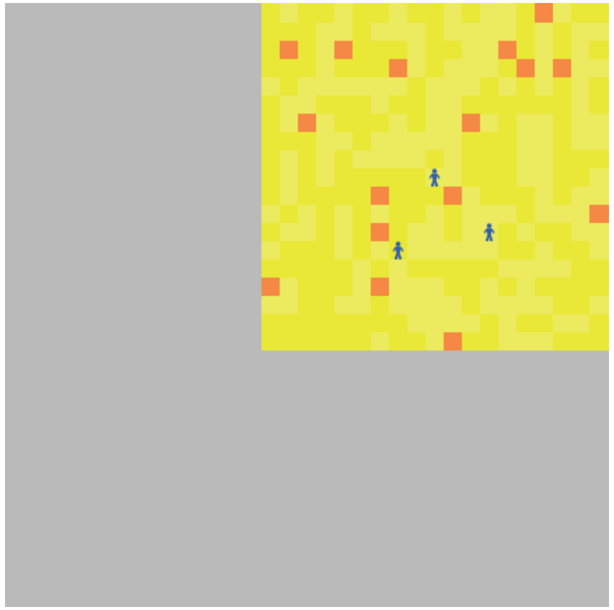
Three-Phased Trial

Pre-trial



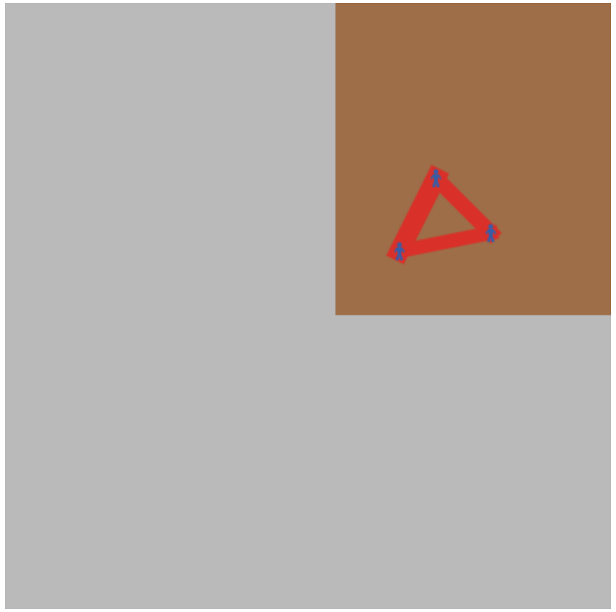
Pre-trial: $S_1>0$, $\omega=0$

Courtroom



Courtroom: $S_2>0$, $\omega=0$

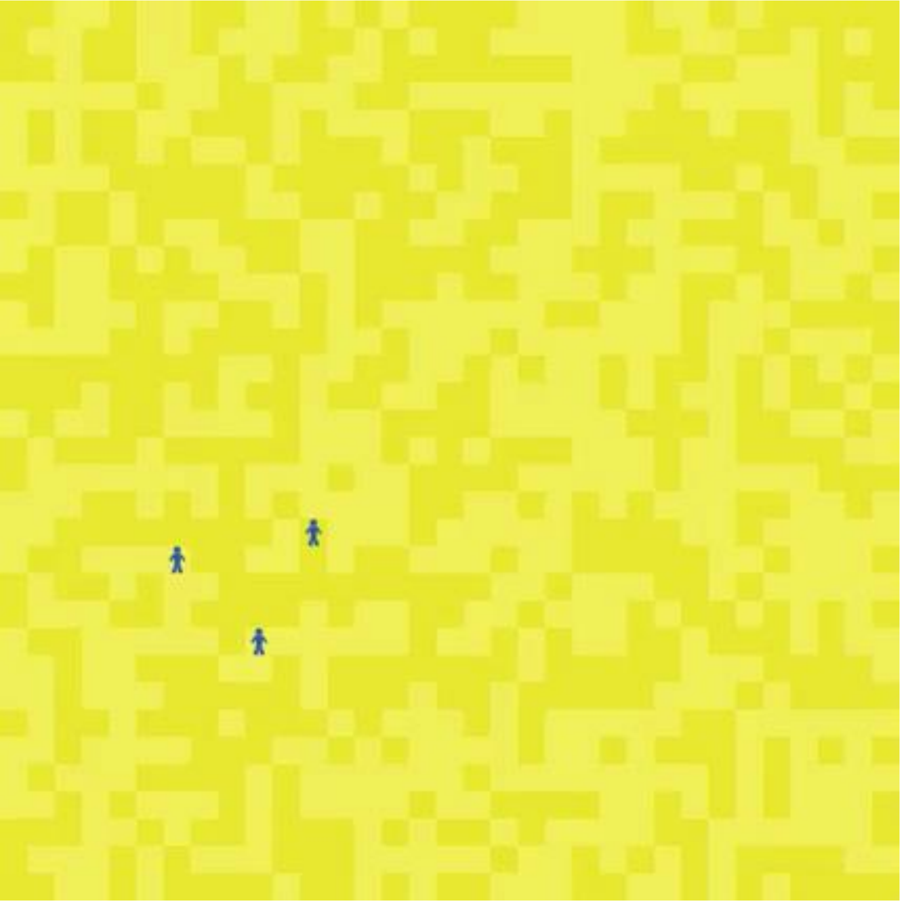
Jury Chamber



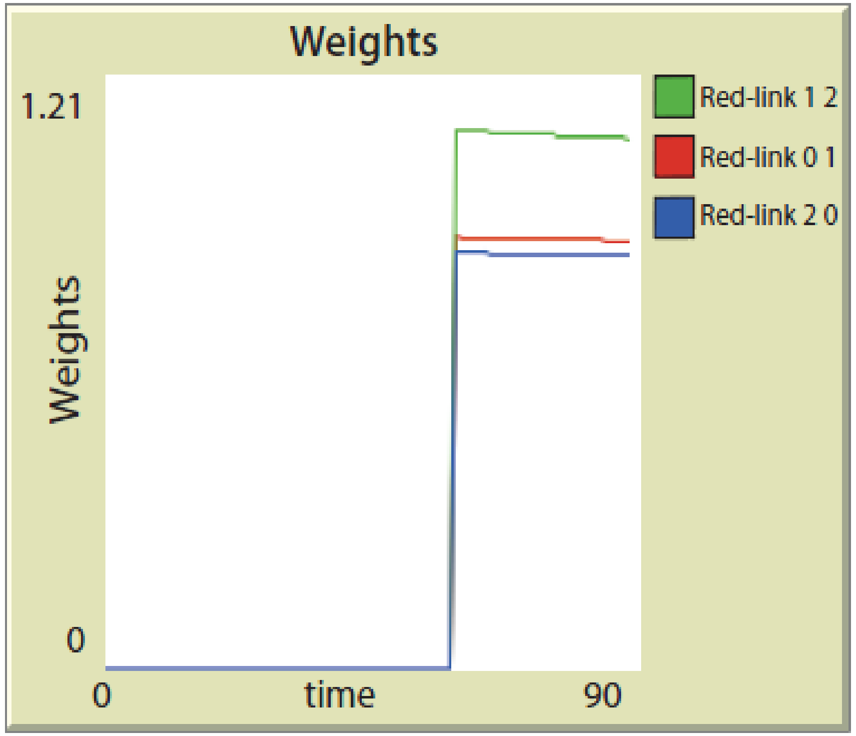
Jury Phase: $S_3=0$, $\omega>0$

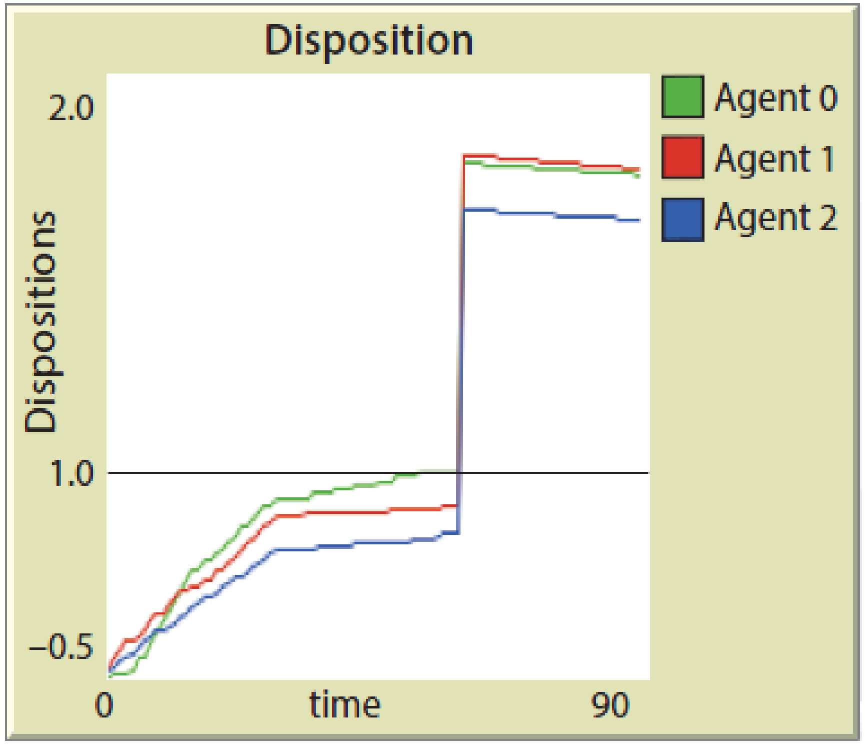
Winston Churchill, “Why, you may take the most gallant soldier, the most intrepid airman, the most audacious soldier, put them at a table…what do you get? The sum of their fears.”

Jury Trial



Weights Jump in Jury Chamber. Drive Dispositions to Convict





Universal Self-Betrayal

$$\forall i, D_i^{total} > \tau > D_i^{solo}$$

***No** jurors would have convicted before the jury phase, but they are unanimous in rendering a guilty verdict, having interacted directly.*

Why Do Networks Happen?

- “Network structure” : links exist or not: {0,1}
- We have continuous weight dynamics but want to study structure.
- How to filter the continuous affinity dynamic onto binary structure of nodes and edges?
- Introduce a link threshold τ_L
- Link exists if and when $\omega_{ji}(t)$ exceeds τ_L

Mathematically

$$L_{ij}(t) = \begin{Bmatrix} 1 \text{ if } \omega_{ij}(t) > \tau_{\mathrm{L}} \\ 0 \text{ otherwise} \end{Bmatrix}.$$

using the Heaviside unit step function

$$L_{ij}(t) = H(\omega_{ij} - \tau_L).$$

Different Thresholds Yield Different Structure History

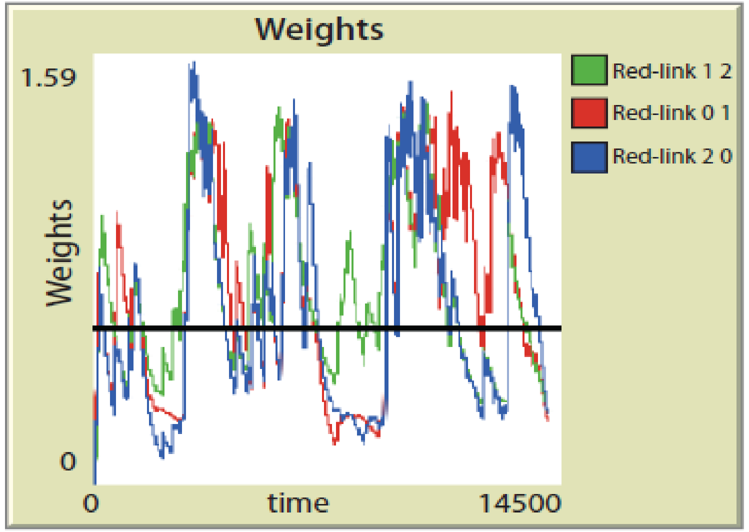
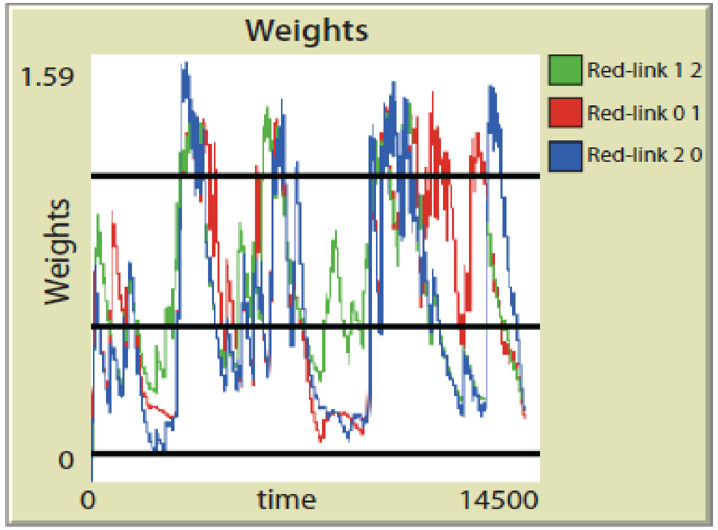


FIGURE 87. Episodes of Connection



Network Structure a “Poincare Map” of Continuum Affinity Dynamic

- Different Poincare Sections (Link thresholds) yield *different structure* dynamics, for *same* affinity dynamic.

Structure as Function of Threshold

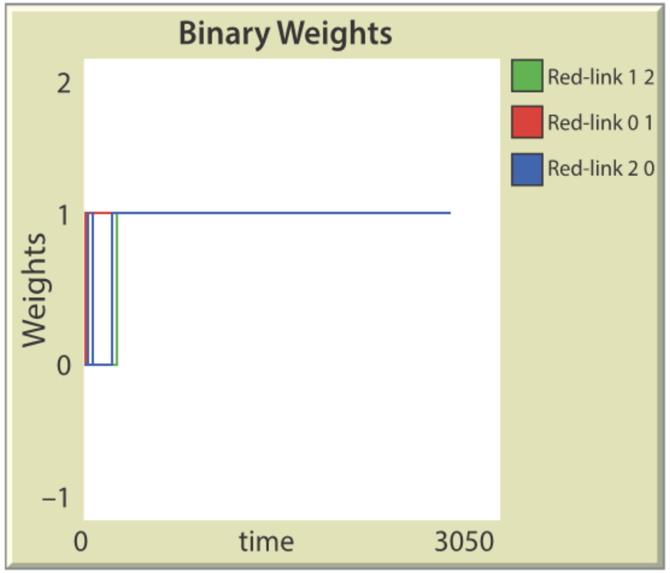


FIGURE 94. Threshold 0.25

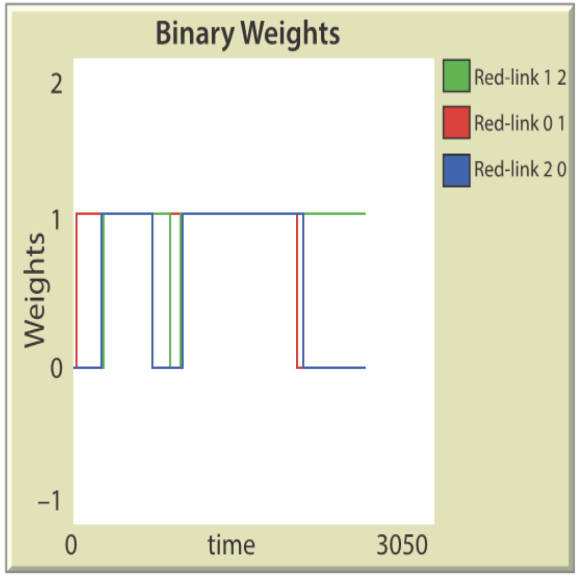


FIGURE 95. Threshold 0.5

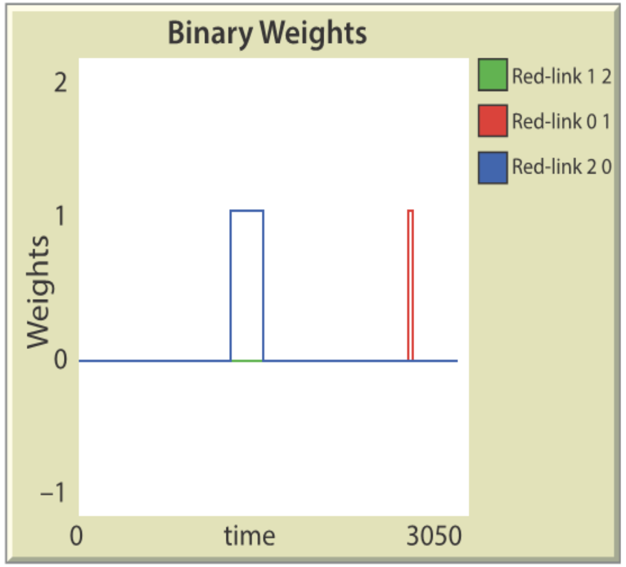
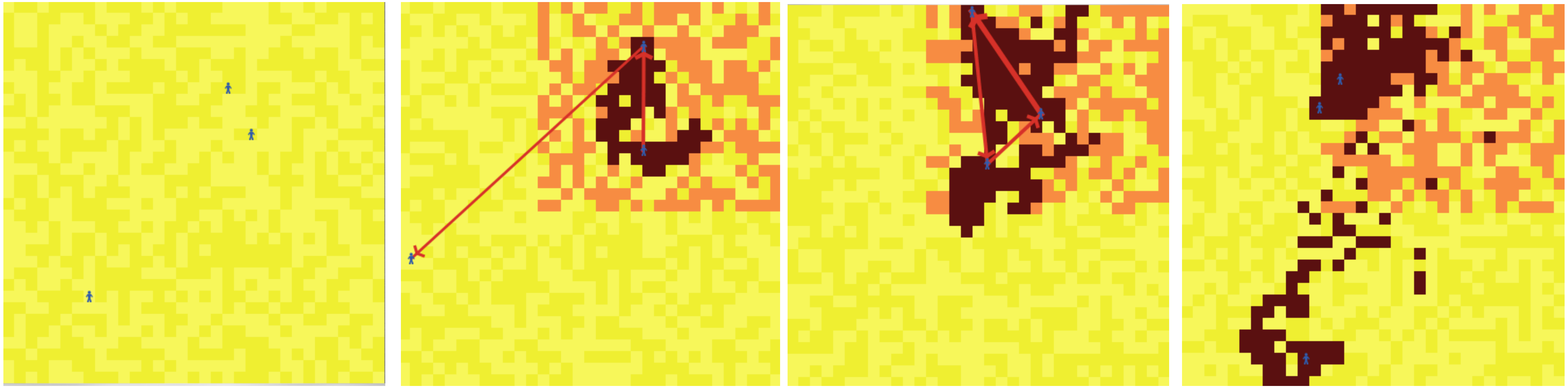


FIGURE 96. Threshold 1.0

For the same affinity dynamics, different thresholds produce different structure dynamics

Link Formation and Breakage: Dynamics of Structure Proper



None Partial Full None

Sexual networks are obviously dynamic and crucial to STD dynamics.

Departure From Literature

- Not preferential attachment. Long mathematical history, moderns (Barabasi and Albert, 1999)
- Rather, Attachment a function of strength and homophily—degree independent attachment
- Lazer 2001(homophily) and Levitan and Wisser 2009 (strength)
- Testable Hpothesis

Part III: Extensions and Foundational issues

Epstein, JM and Chelen J , “Advancing Agent Zero” in Kirman A and Wilson DS, eds. *Complexity and Evolution: Toward a New Synthesis for Economics* (MIT Press, 2016)

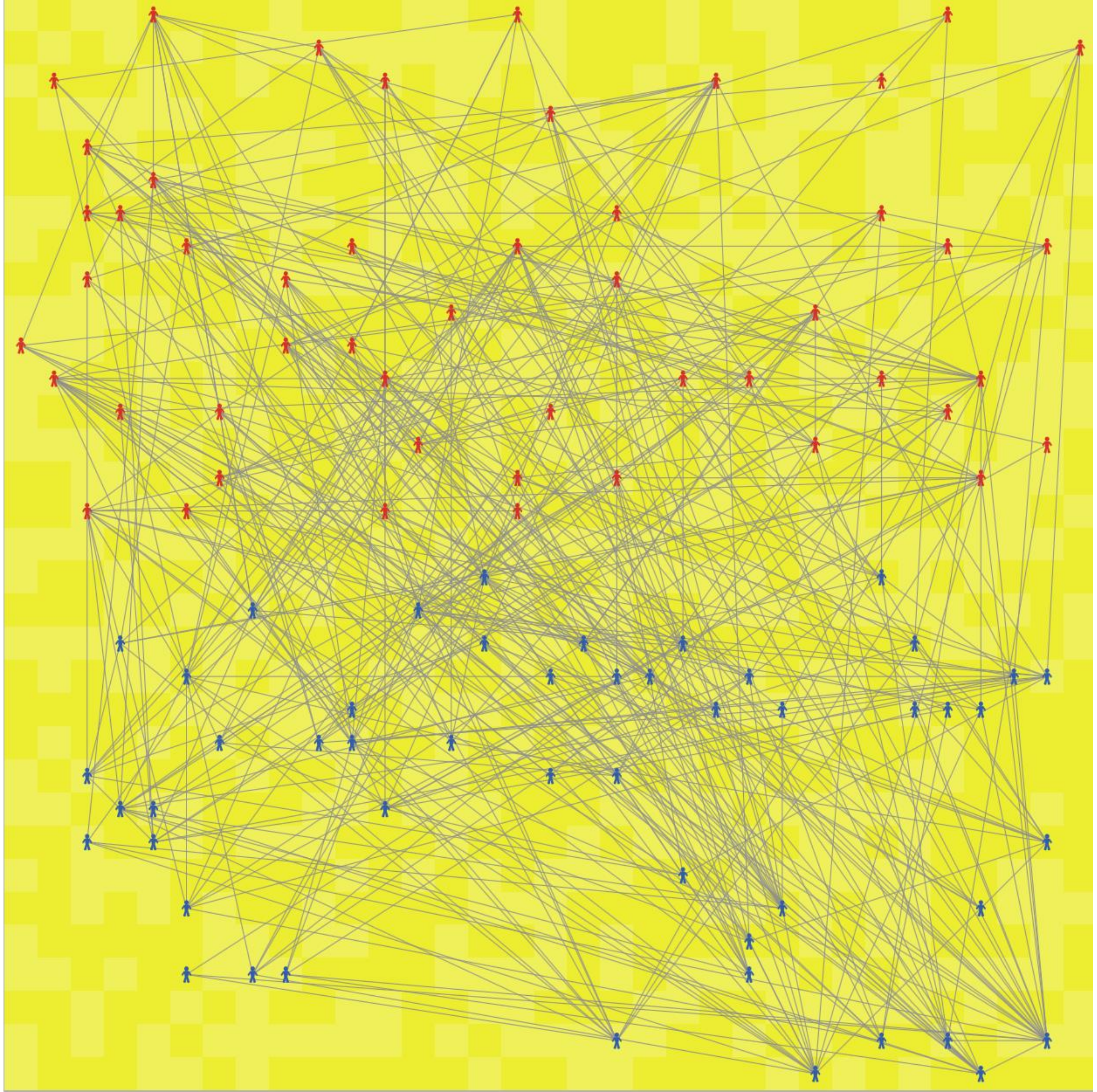
Extensions

- Scale-Up to large numbers
- Permit arbitrary network topologies
- The most arresting Parable, to me, is the first actor:
- An agent that has no aversive stimulus and would not act alone leads the lynch mob, by dispositional contagion.
- How robust is that?
- For arbitrary numbers of northerners (with stimulus) and southerners (without) and arbitrary network topologies, it is mathematically formidable.
- Preliminary computational experiments very interesting.

Explorations

Can scale up and stipulate fixed network structures and explore dynamics computationally

Exponential degree distribution (λ=5)

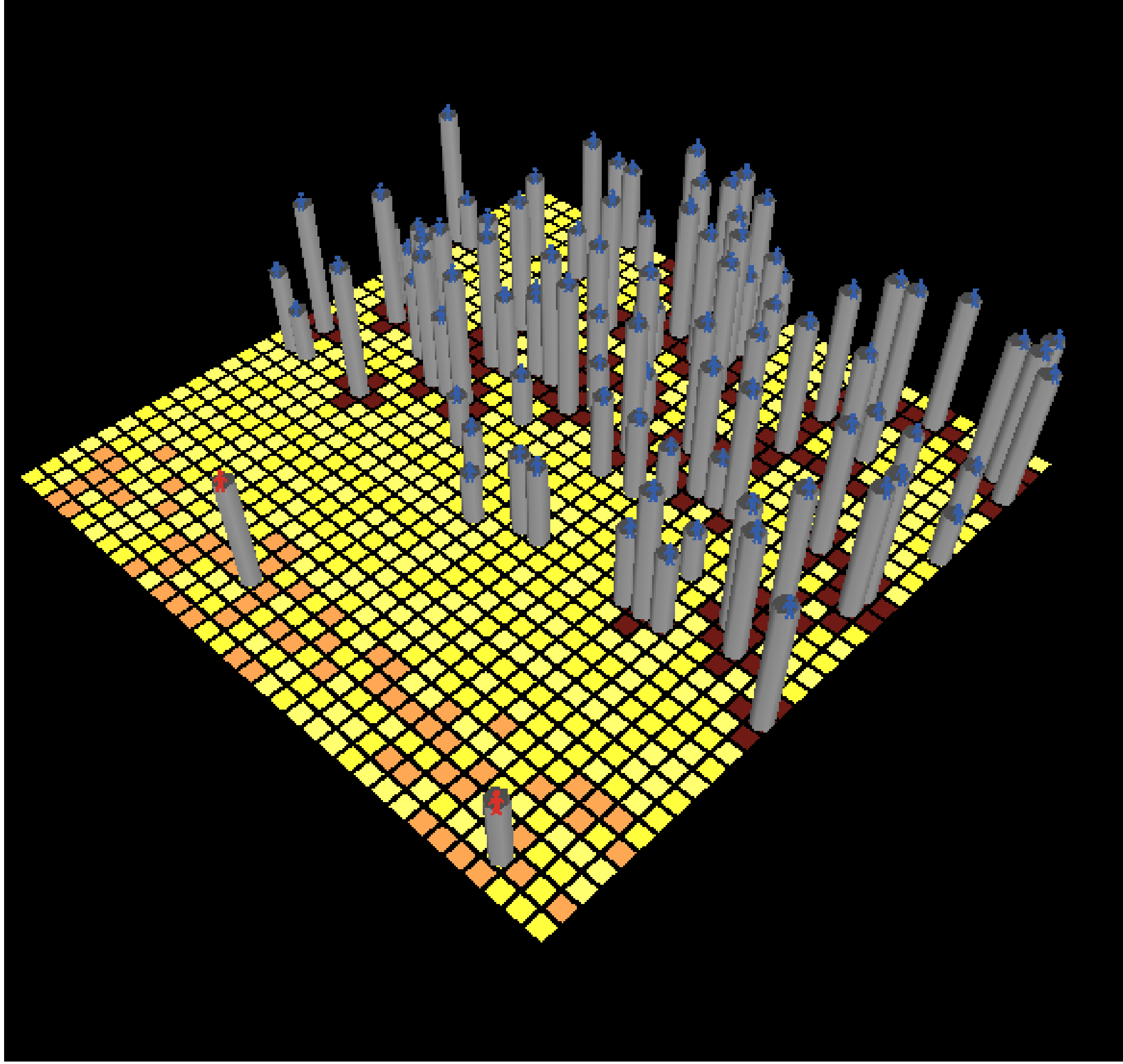


Turn off all the movies,
assume distributions
and prove some theorems
on core phenomena:

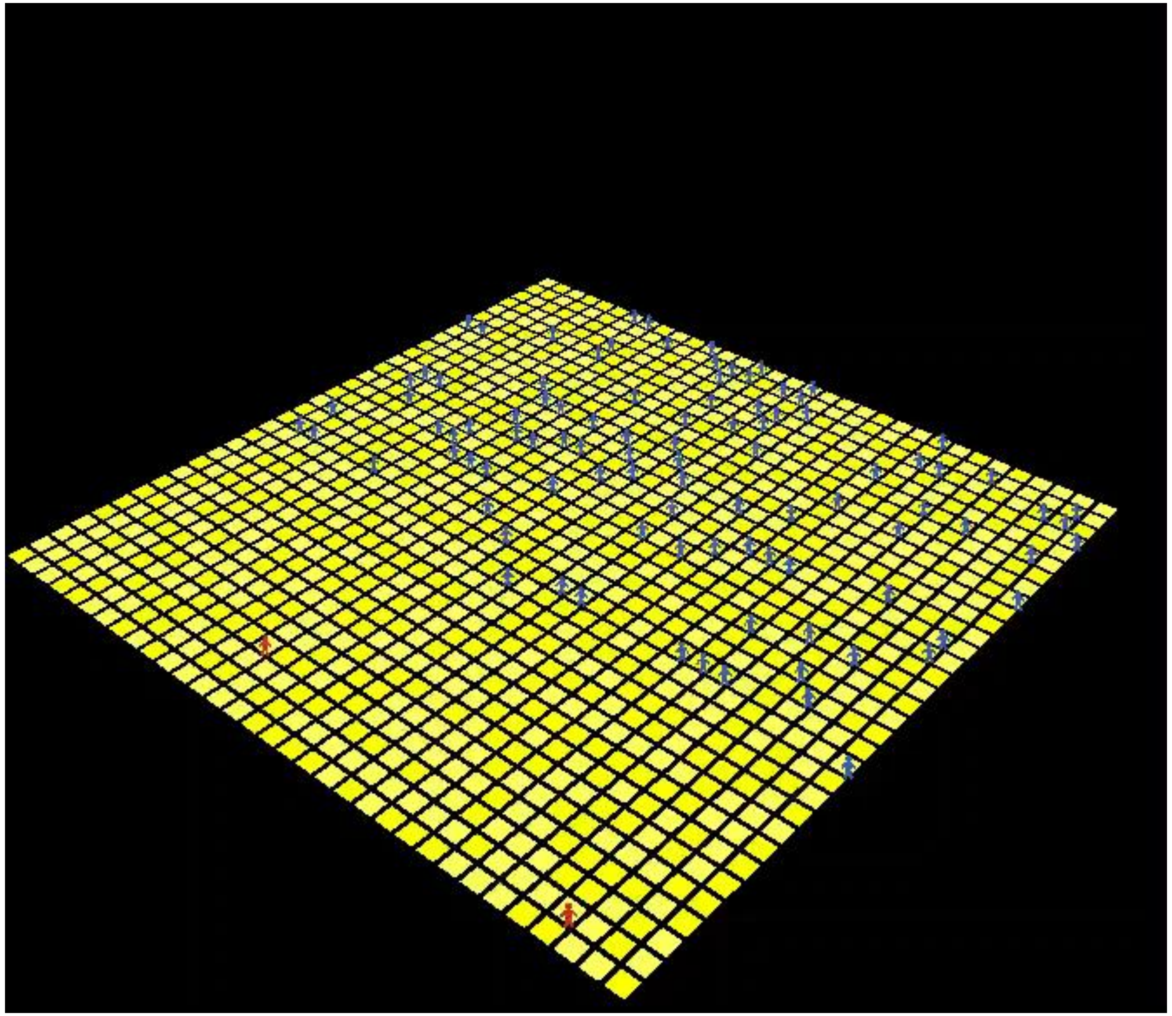
[1] waiting time to first actor,

[2] probability of universal self betrayal.

Large-Scale Activation without direct stimulus by Dispositional Contagion



Large-Scale Activation without Direct Stimulus



Toward a Theory of the First Actor

- Our Post-Doc Jeewoen Shin and I have some preliminary analytics
- And are pursuing this computationally
- But it is clearly more than an outlier of significant interest to the study of cascading social phenomena.

Inverse Generative Social Science

- Machine learning is augmenting humans, crushing them at chess, and replacing them, but it is not *explaining* them! It can be used to do that.
- Using evolutionary programming to discover ABMs that generate target macro-data.
- Typically hand-crafted, including *Agent_Zero*.
- Stipulate only minimal code components and concatenation operators (mathematical, logical) and evolve fittest AMBs.
- Data-driven evolution of generative micro-models
- "Toward Inverse Generative Social Science using Multi-Objective Genetic Programming" Vu, Probst, Epstein, et al. *GECCO* 2019.
- Founding Workshop in January 2019 Washington, DC.

- Thank You!

- je65@nyu.edu