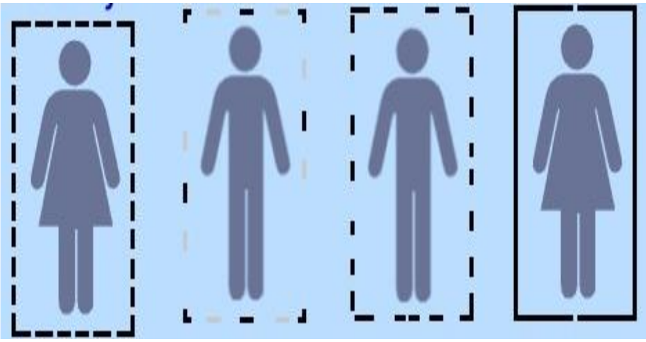
UMassAmherst

LIMINALITY AND DIFUSSION IN SOCIAL NETWORKS: AN AGENT-BASED COMPUTATIONAL MODEL

Diego F. Leal
Dept. of Sociology
University of Massachusetts, Amherst
(www.diegoleal.info)

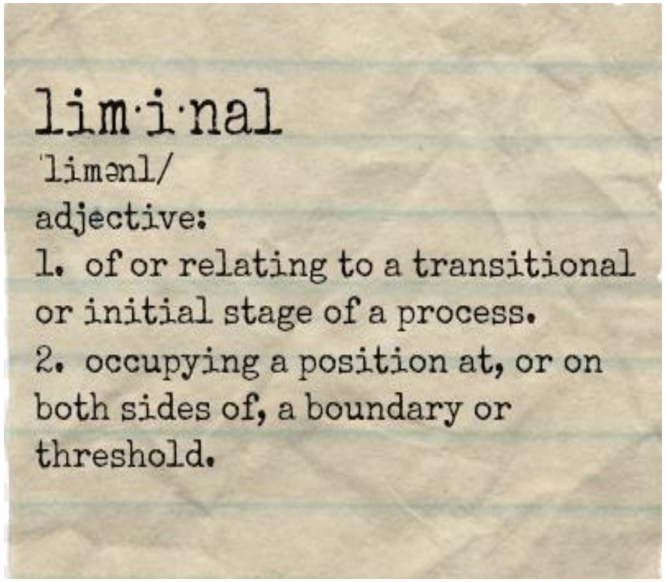
The Big Picture: Symbolic Boundaries



- A primary interest for sociologists
  (Lamont and Molnár 2002)
- Dynamic and permeable
  (Pachucki et al 2007; Mäs et al.2014)
- Often intersect
  (Browne and Misra 2003)

Liminality

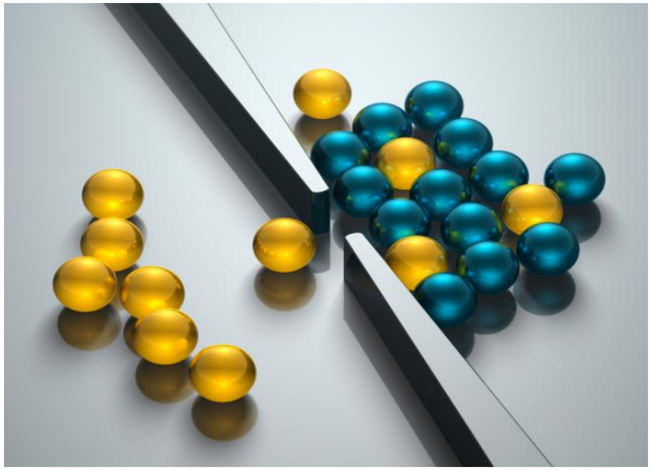
“*The state of being associated with people who are simultaneously members of two or more culturally distinct groups, which allow them to move beyond an “either/or” to a “both/neither” path of identification*?” (Romo 2011)



Liminal agents are by definition culturally liminal

Liminality across Which Boundary?

Bright boundaries
(Alba 2005)



Liminality across Which Boundary?

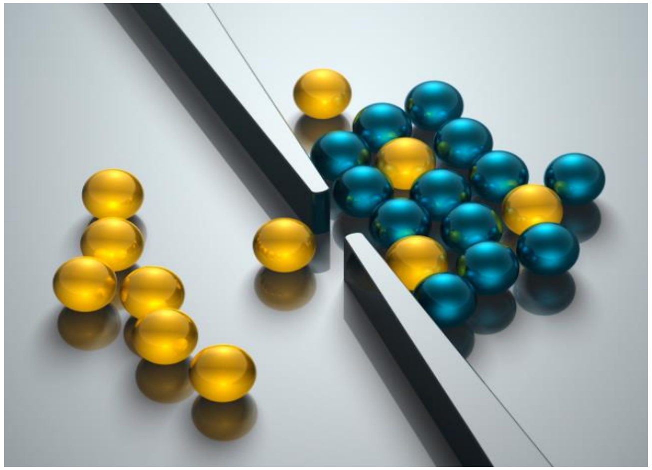
Bright boundaries
(Alba 2005)

↓

Ethnoracial boundary

↓

Segregation



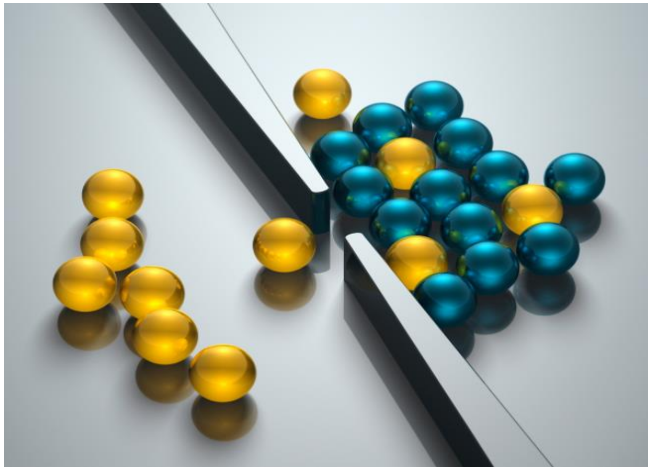
Liminality across Which Boundary?

Bright boundaries
(Alba 2005)

Ethnoracial boundary

Segregation

Cultural holes



Ethnoracial Boundaries and (Ethno)Racial Brokers

I expect to find positive evidence that liminal agents are culturally <u>skilled</u> to perform the <u>difficult</u> task of bridging cultural holes (i.e. reducing segregation in networks)

The Agent-Based Computational Model

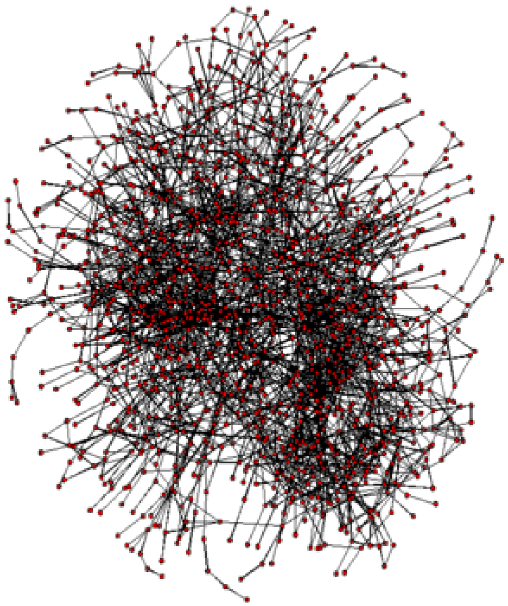
Populate the Model: Pseudo Code

- Extract Add Health friendship nomination network data (e.g. Sunshine High data).
- Create as many agents as there are in the school under analysis (n)
- Bestow each agent with the attributes (sex, race, grade, income, and ethnicity) of the student it represents.
- Create a variable for each agent, A_i, that signals adoption (A_i = 1 adopted; A_i =0 did not adopt)

Populate the Model: Describing Sunshine High

Network Descriptive Statistics		Race & Ethnicity	%
Number of nodes	1,452	White	18.18
Number of dyads	2,979	Black	19.15
Isolates*	193 (11.2%)	Native American	1.79
Non-isolates out largest comp.*	31 (1.8%)	Asian	33.61
Graph density	0.003	Other	21.42
Mean geodesic distance	1.37	Multiracial	5.85
Mean degree	4.10	Latino (any race)	39.94
Modal degree	2		
Avg. local clustering coeff.	0.18		
Mean path length	6.49		
Degree centralization	0.01		
Ln[Gross friendship. Segreg.]**	1.63		
Size largest clique	6		

Sociogram (Largest Component)



* Before retaining the largest component.
**Gross friendship segregation (α) "is substantively interpretable as the odds ratio of a friendship between members of a same-race dyad relative to friendship in a cross-race dyad. When alpha = 1, then the odds of a same-race friendship equal the odds of a crossrace friendship, and the setting is perfectly integrated. As alpha increases, the relative odds of a same-race friendship increase by a factor of alpha. Since α is scaled from 0 to infinity, I use ln(α), which ranges from - infinity to infinity." Moody (2001: 692)

Modeling Diffusion: Pseudo Code

- For agent i to agent n, where n is the number of agents:
  - Make agent i the seed innovator ($I_i = 1$)
  - Make i and i's group of friends adopt the innovation ($A_i = 1$).
  - Make the members of i's egonet be part of the seed neighborhood ($S_i = 1$).

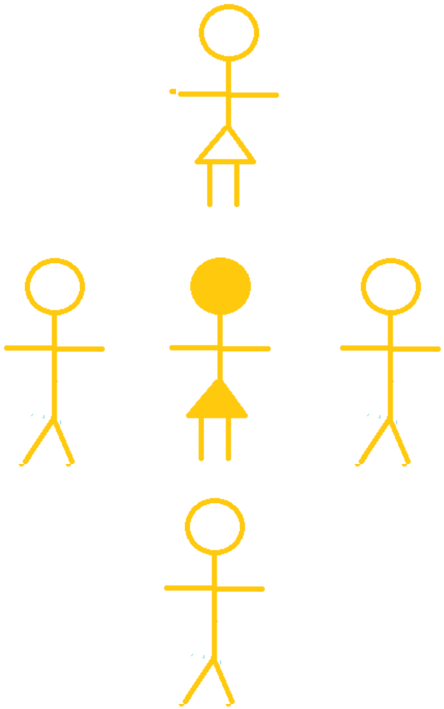
Modeling Diffusion: Assigning $I_i = 1$

Filled = Adopted
Not Filled = Did not Adopt

The seed innovator ($I_i = 1$)

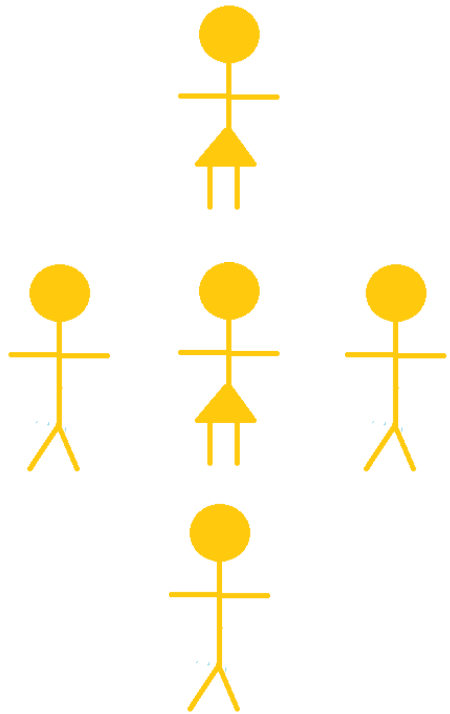
Modeling Diffusion: Identify i's neighbors

Filled = Adopted
Not Filled = Did not Adopt



Modeling Diffusion: Assigning $S_i = 1$

Filled = Adopted
Not Filled = Did not Adopt



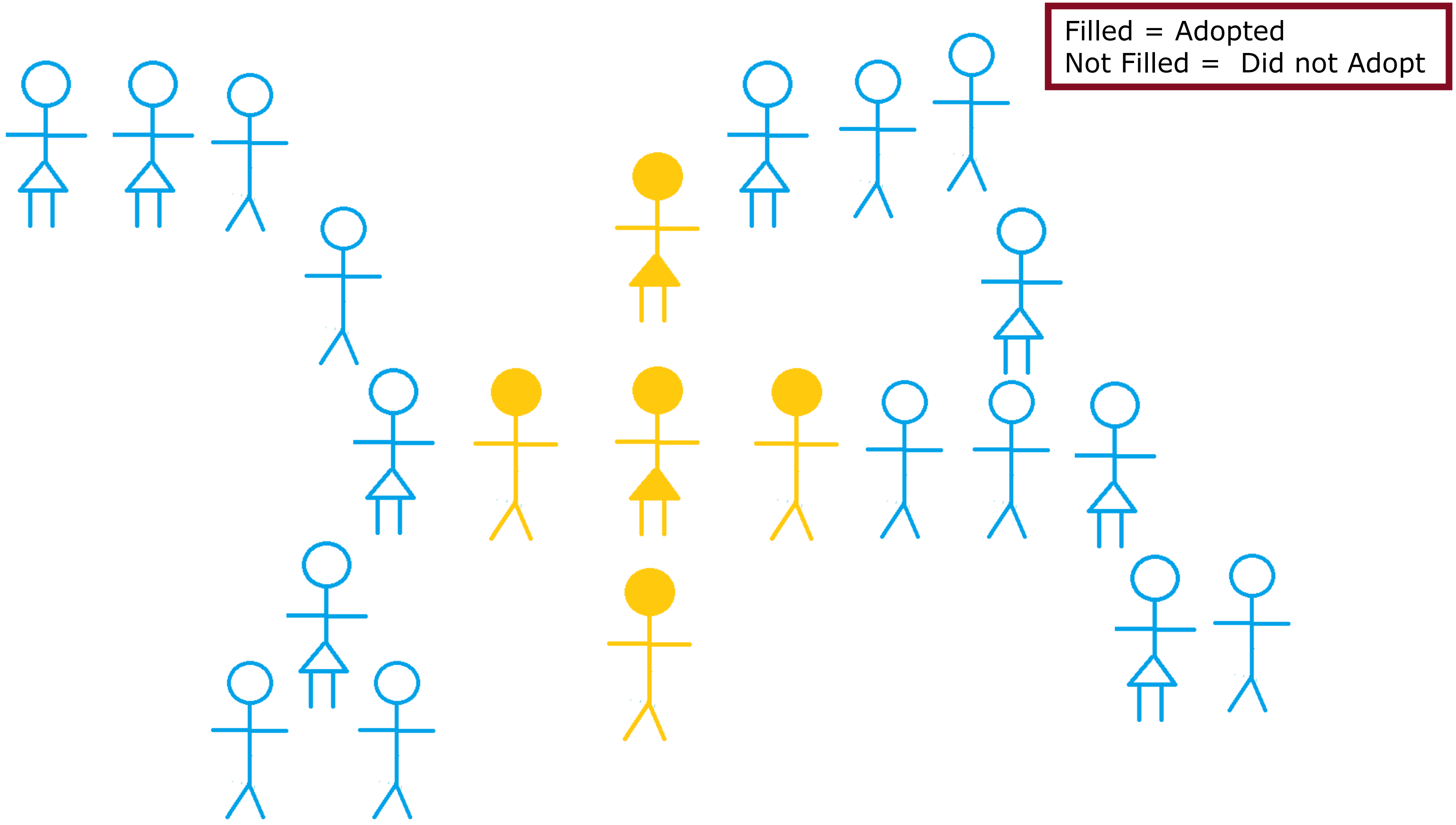
The seed neighborhood. All these agents have $S_i = 1$ & $A_i = 1$

Modeling Diffusion: Pseudo Code

- For agent i to agent n, where n is the number of agents:
  - Make agent i the seed innovator ($I_i = 1$)
  - Make i and i's group of friends adopt the innovation ($A_i = 1$).
  - Make the members of i's egonet be part of the seed neighborhood ($S_i = 1$).

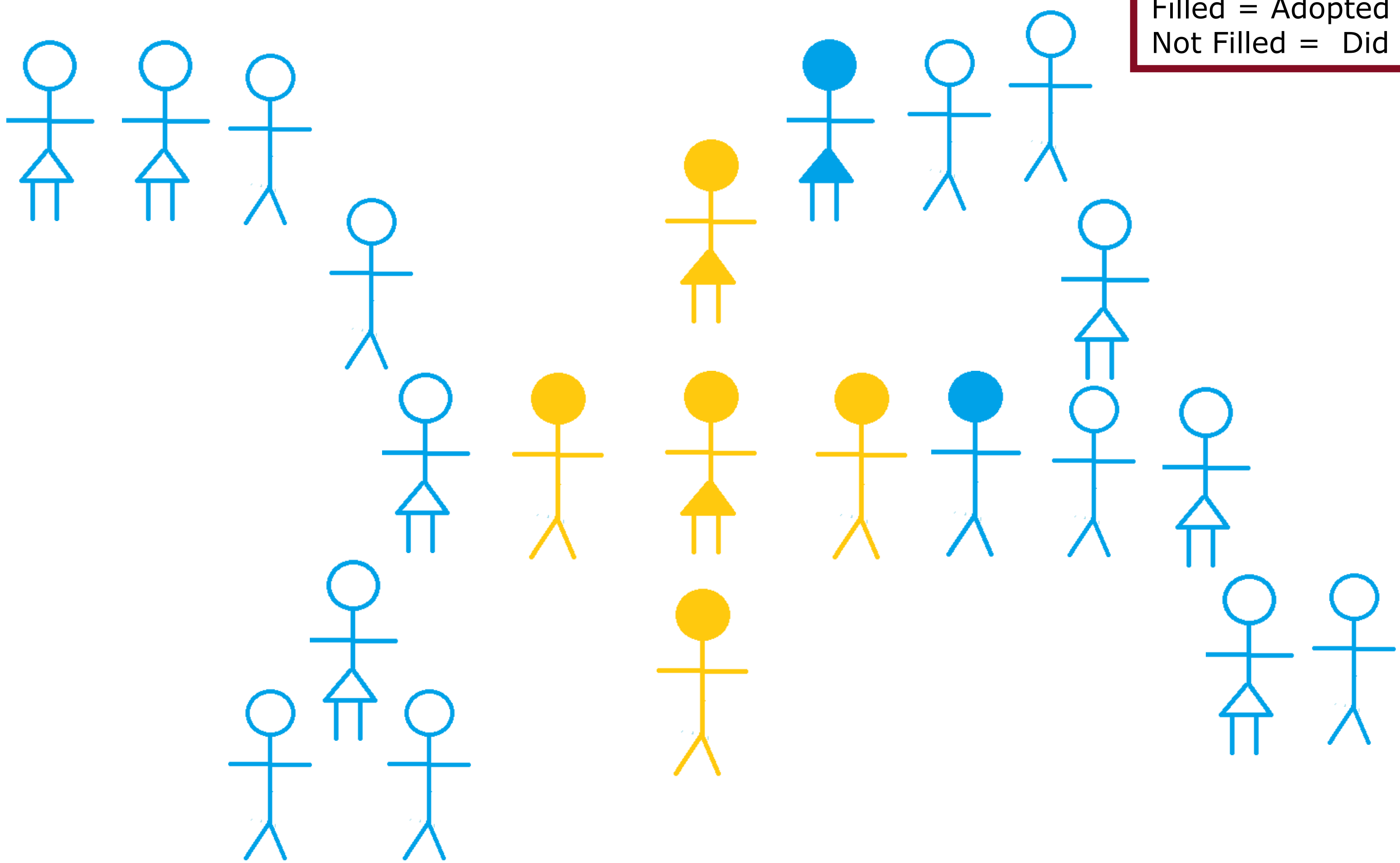
  - Repeat R times ($R \in [1,n]$) in random order, with no replacement:
    - Ask agents with $S_i = 0$ to:
      - Adopt the innovation ($A_i = 1$) if a given user-defined proportion, T, of her friends have adopted the innovation ($T \in [0,1]$)

Modeling Diffusion: Threshold Effect, R = 0

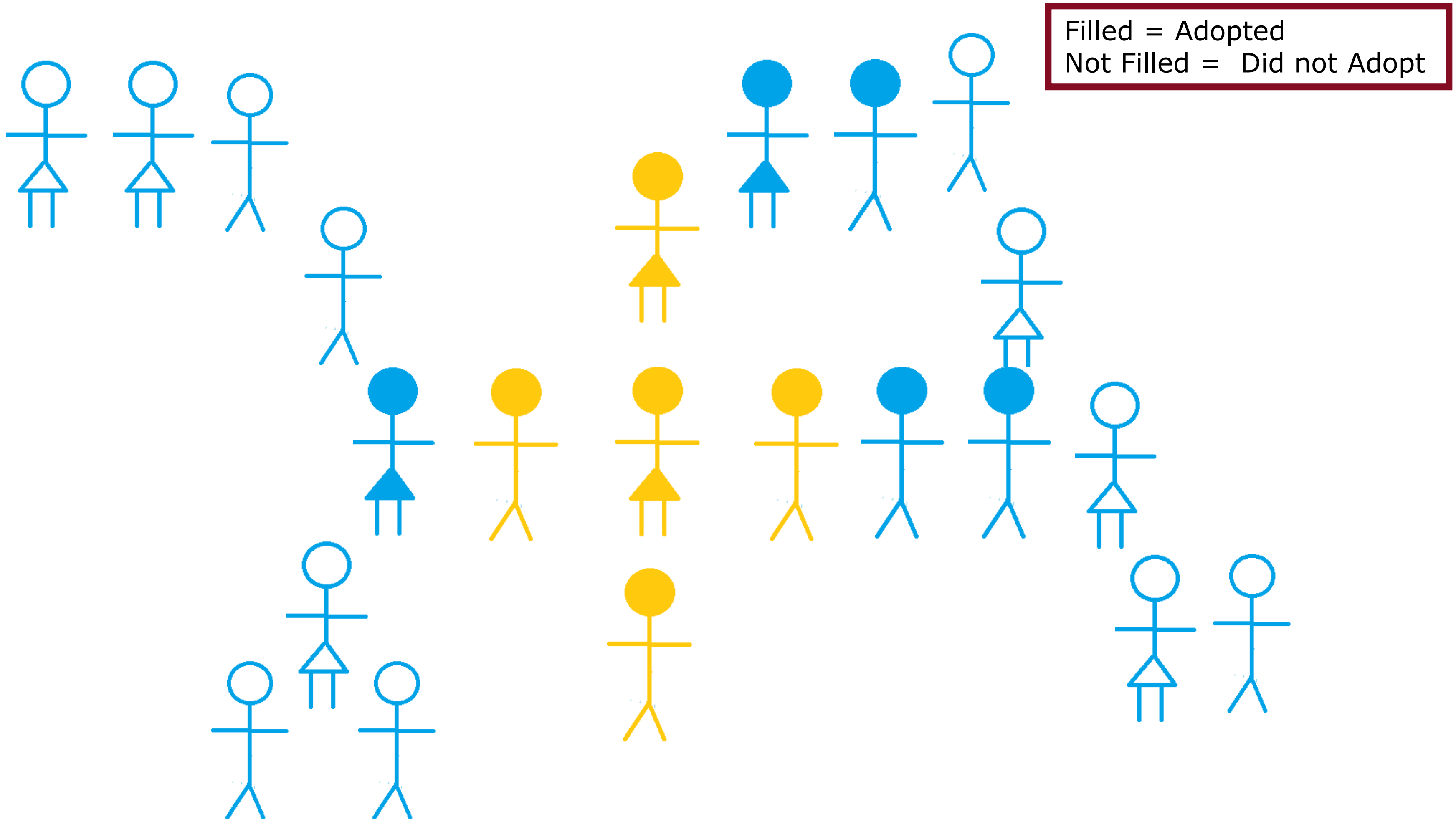


Modeling Diffusion: Threshold Effect, R = 1

Filled = Adopted
Not Filled = Did not Adopt

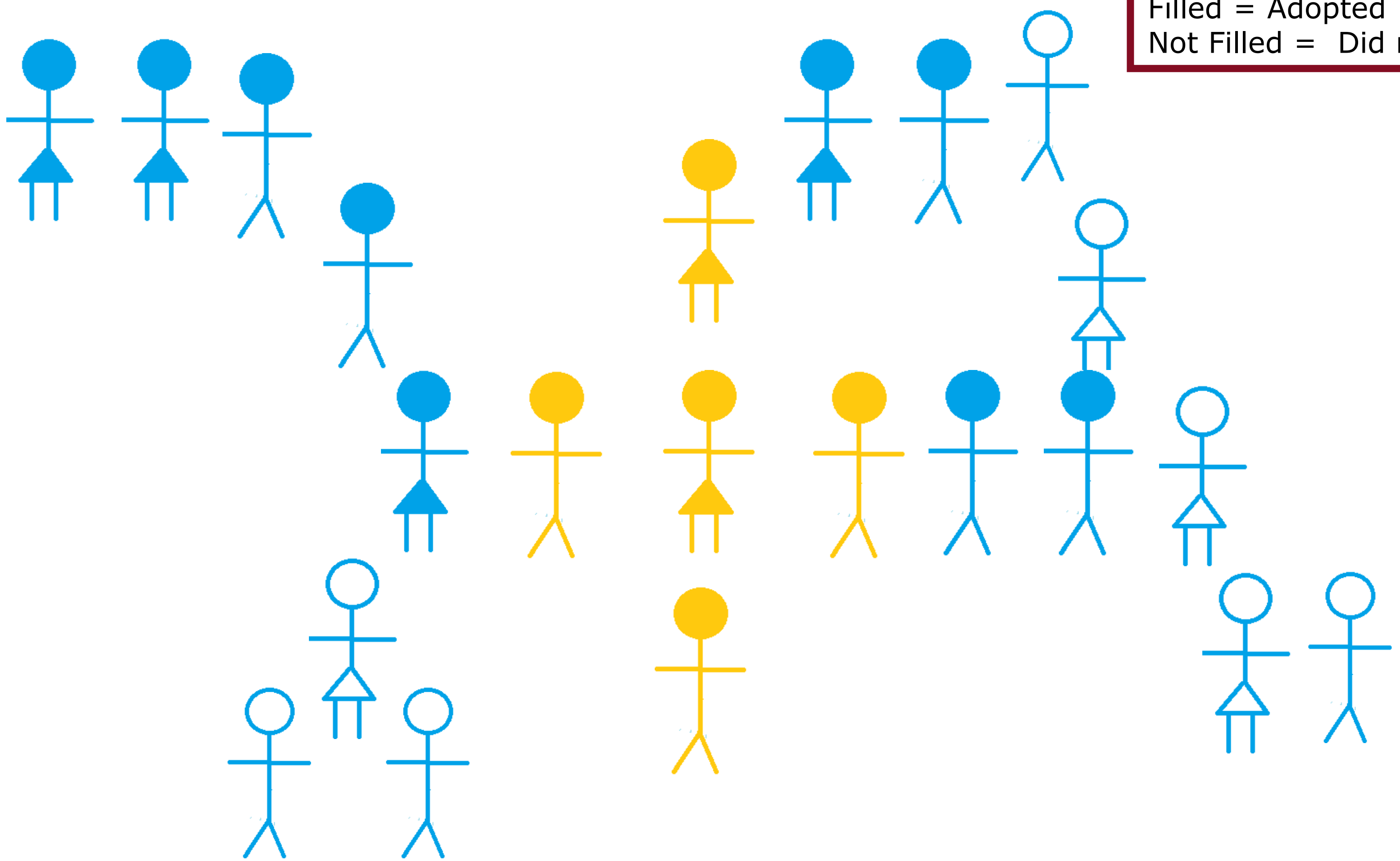


Modeling Diffusion: Threshold Effect, R = 3



Modeling Diffusion: Threshold Effect, R = 4

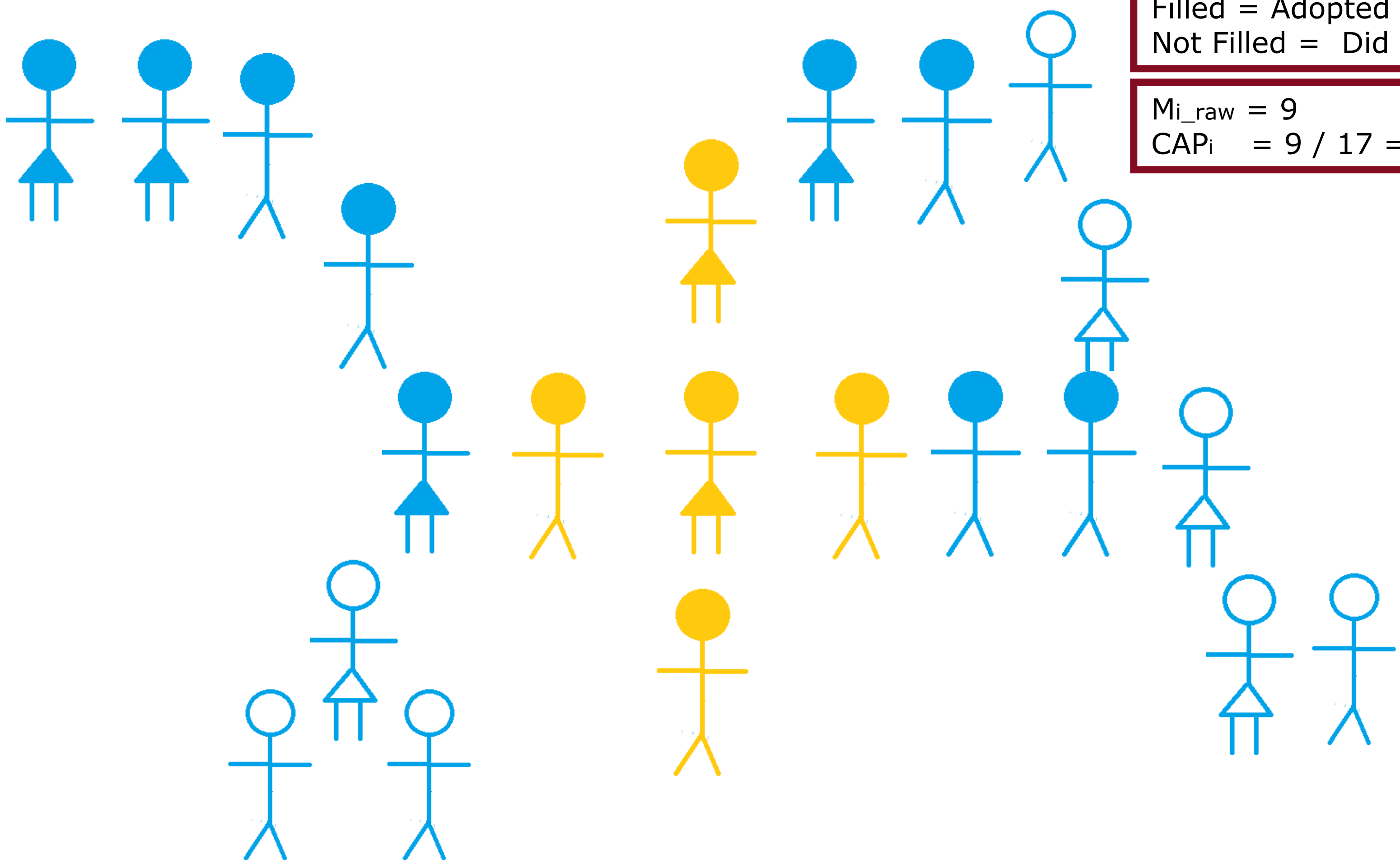
Filled = Adopted
Not Filled = Did not Adopt



Modeling Diffusion: Pseudo Code

- For agent i to agent n, where n is the number of agents:
  - Make agent i the seed innovator (I_i = 1)
  - Make i and i's group of friends adopt the innovation (A_i = 1).
  - Make the members of i's egonet be part of the seed neighborhood (S_i = 1).
  - Repeat R times (R ∈ [1,n]) in random order, with no replacement:
    - Ask each agent with S_i = 0 to:
      - Adopt the innovation (A_i = 1) if a given user-defined proportion, T, of her friends have adopted the innovation (T ∈ [0,1])
  - Retrieve the number of agents infected when agent i is the seed innovator (Σ Ai - Σ S_i). Call this quantity M_{i_raw}
  - Calculate the proportion of the population infected by agent i (M_{i_raw} / (n – (Σ S_i))). This is the *spreading capacity* of agent i (CAP_i).

Modeling Diffusion: i's Spreading Capacity (CAP_i)



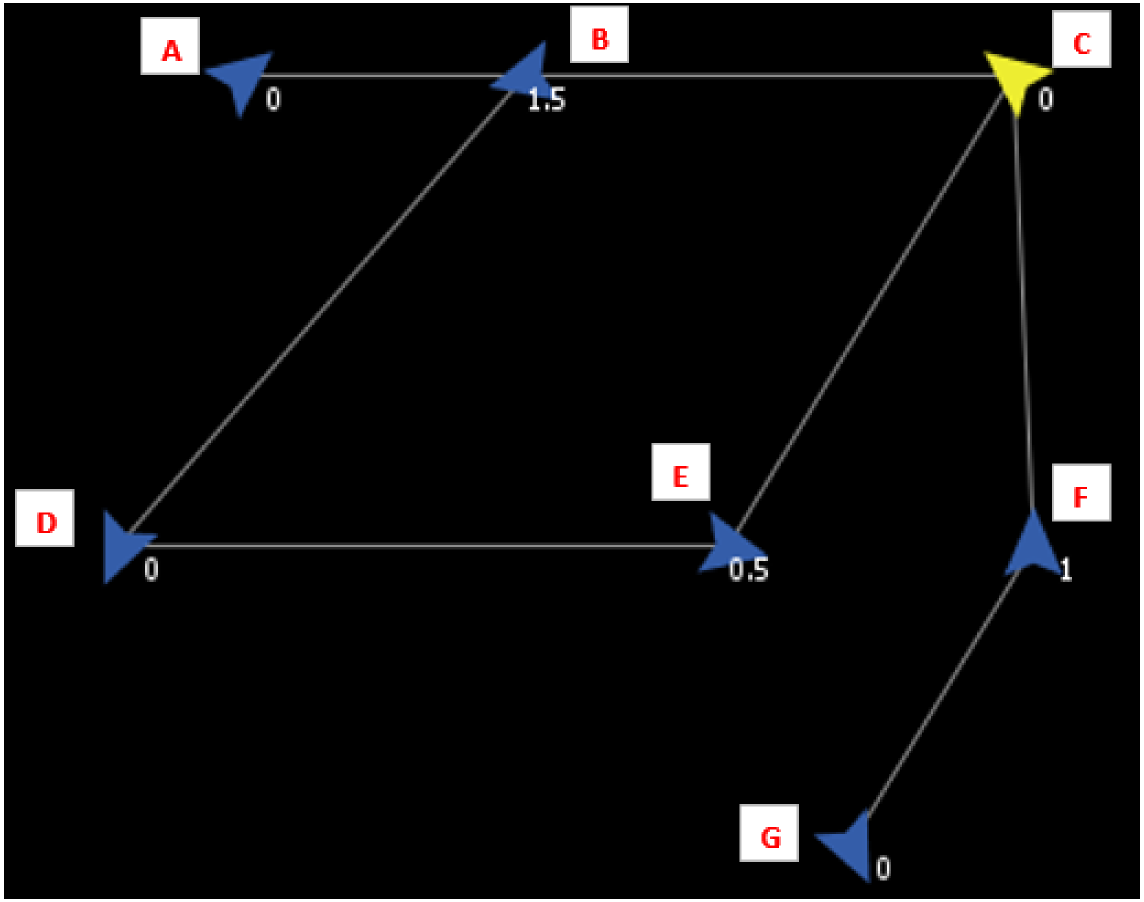
Filled = Adopted
Not Filled = Did not Adopt

M_{i_raw} = 9
CAP_i = 9 / 17 = 0.53

Measuring Liminality: Pseudo Code

- Calculate liminality scores for each ego:
  - Betweenness ($betw_i$); interracial brokerage (IB_i); egonet diversity ($diver_i$); egonetwork diversity * ego's degree ($diverDeg_i$)

Measuring Liminality : Capacity for IR Brokerage



$$IB_j = \sum_{i=1;i<j}^{n} \sum_{k=1}^{n} \frac{b_{(ijk)}}{p_{(ik)}} * m_{(ik)}$$

Where $b_{(ijk)}$ is equal to 1 if actor i is connected to j, j is connected to actor k, and actor i is not connected to actor k, and 0 otherwise; $p_{(ik)}$ is the number of 2-step undirected paths between i and k, and $m_{(ik)}$ equals 1 if i and k are from different races (or ethnicities, if the focus is on interethnic brokerage), and 0 otherwise.

Based on: Gould (1988)

Extends: Pachucki and Breiger (2010)

Measuring Diffusion: Pseudo Code

- Calculate liminality scores for each ego:
  - Betweenness ($betw_i$); interracial brokerage (IB_i); egonetwork diversity ($diver_i$); egonetwork diversity * ego's degree ($diverDeg_i$)
- Consider a user-defined proportion *p* of the most efficient spreaders ($p \in [0,1]$)
- Include in the set $\Upsilon_{CAP}(p)$ the portion *p* of agents with the highest spreading capacity (CAP_i).
- Compute the average CAP_i for the egos included in $\Upsilon_{CAP}(p)$. Call this quantity $AVG_{CAP}(p)$. **This is the *de facto* highest average spreading capacity in the system, as function of p.**

Modeling Diffusion

Agent ID	CAP
017	1
921	1
432	0.99
008	0.76
234	0.67
21	0.67
465	0.66
92	0.65
231	0.65
AVG_{CAP}	0.80

Assume:
n = 90 ;given by the data
p = 0.10 ;user defined

Then we have:
Size of $\Upsilon_{CAP} = 9$

Measuring Diffusion: Pseudo Code

- Calculate liminality scores for each ego:
  - Betweenness ($betw_i$); interracial brokerage (IB_i); egonet diversity ($diver_i$)
- Consider a user-defined proportion p of the most efficient spreaders ($p \in [0,1]$)
- Define the set $\Upsilon_{CAP}(p)$ as the portion p of agents with the highest spreading capacity (CAP_i).
- Compute the average CAP_i for the egos included in $\Upsilon_{CAP}(p)$. Call this quantity $AVG_{CAP}(p)$. **This is the *de facto* highest average highest spreading capacity in the system, as function of p.**
- Define the set $\Upsilon_{liminality\_measure}(p)$ as the portion p of agents with the highest score in a given liminality measure. For instance, the set $\Upsilon_{betw}(p)$ is the portion p of agents with the highest betw centrality.
- Compute the average proportion of alters infected for each ego included in $\Upsilon_{liminality\_measure}(p)$. Call this quantity $AVG_{CAP\_liminality\_measure}(p)$

Modeling Diffusion

Agent ID	CAP	Agent ID	CAP_{betw}	Agent ID	CAP_{IR_Broker}
017	1	1342	0.88	876	0.95
921	1	121	0.76	534	0.90
432	0.99	43	0.76	456	0.89
008	0.76	543	0.73	565	0.75
234	0.67	1231	0.72	213	0.76
21	0.67	32	0.56	002	0.60
465	0.66	76	0.56	345	0.59
92	0.65	987	0.54	543	0.59
231	0.65	687	0.53	765	0.59
AVG_{CAP}	0.80	**AVG_{betw}**	0.69	**AVG_{IR_Broker}**	0.73

Assume:
n = 90 ;given by the data
p = 0.10 ;user defined

Then we have:
Size of $\Upsilon_{CAP} = \Upsilon_{CAPbetw} = \Upsilon_{CAP\_IR\_broker} = 9$

Measuring Diffusion: The Imprecision Function

- Following Kitsak et al. (2010), I compute an *imprecision function* (IM) based on these two quantities:

$$IM_{liminality\_measure}(p) = 1 - [AVG_{CAP\_liminality\_measure}(p)/AVG_{CAP}(p)]$$

Modeling Diffusion

Agent ID	CAP	Agent ID	CAP_{betw}	Agent ID	CAP_{IR_Broker}
017	1	1342	0.88	876	0.95
921	1	121	0.76	534	0.90
432	0.99	43	0.76	456	0.89
008	0.76	543	0.73	565	0.75
234	0.67	1231	0.72	213	0.76
21	0.67	32	0.56	002	0.60
465	0.66	76	0.56	345	0.59
92	0.65	987	0.54	543	0.59
231	0.65	687	0.53	765	0.59
AVG_{CAP}	0.80	**AVG_{betw}**	0.69	**AVG_{IR_Broker}**	0.73
		IM_{betw}	0.14	**IM_{IR_Broker}**	0.09

Assume:
n = 90 ;given by the data
p = 0.10 ;user defined

Then we have:
Size of $\Upsilon_{CAP} = \Upsilon_{CAPbetw} = \Upsilon_{CAPir\_broker} = 9$

Results: Imprecision Function (T = 0.25; R = 10)

	Proportion (p) of Top Spreaders			
Imprecision Function	**5%**	**10%**	**15%**	**20%**
Random	0.629	0.625	0.609	0.599
Egonet racial heterogeneity	0.736	0.740	0.745	0.709
Egonet racial het. * degree	0.201	0.318	0.359	0.372
Betweenness centrality	0.168	0.201	0.235	0.266
Capacity for racial Brokerage	0.164	0.197	0.222	0.253

Based on 50 simulations.

When seed innovators are chosen at random, the selection of influential spreaders is ~62% inaccurate (vis-à-vis choosing the *de facto* best spreaders). On the other hand, choosing agents based on their capacity for **interracial brokerage is ~20% inaccurate**, a big improvement vis-à-vis randomness.

Results: Predicting the IR Capacity for Brokerage

	Linear Reg. Coeff. (Std. Error)	Negative Binomial Reg. Coeffi (Std. Error)	Logistic Reg. Odds Ratio (Std. Error)
9th grad (ref)			
10th grade	0.111 (0.366)	-0.138 (0.099)	0.729 (0.138)
11th grade	0.475 (0.486)	-0.180 (0.128)	0.898 (0.224)
12th grade	0.430 (2.378)	0.151 (0.650)	3.525 (4.332)
Men (ref)			
Women	0.230 (0.248)	0.043 (0.066)	1.031 (0.133)
Low SES (ref)			
Middle SES	-0.478 (0.352)	-0.039 (0.093)	0.829 (0.152)
High SES	0.069 (0.316)	0.040 (0.084)	0.930 (0.154)
14 years old (ref)			
15 years old	-0.188 (0.432)	-0.112 (0.115)	0.654 (0.148)
16 years old	-0.114 (0.518)	-0.180 (0.137)	0.692 (0.186)
17 years old	0.286 (0.617)	-0.287† (0.164)	0.530 (0.171)
18 years old	0.280 (0.876)	-0.314 (0.240)	0.654 (0.292)
19 years old	-3.739 (2.874)	-1.132 (0.769)	0.083† (0.111)
White (ref)			
Asian	-2.556*** (0.490)	-0.792*** (0.126)	0.158*** (0.039)
Black	-0.467 (0.519)	-0.378** (0.137)	0.367*** (0.092)
Native	0.210 (0.747)	-0.197 (0.190)	0.748 (0.292)
Other	-0.356 (0.395)	-0.012 (0.098)	0.818 (0.169)
Multiracial	1.393** (0.596)	0.622*** (0.154)	3.263*** (1.045)
Latino (any race)	1.220* (0.485)	0.212† (0.123)	1.868** (0.450)
Degree	1.835 *** (0.048)	0.482*** (0.015)	1.754*** (0.064)
Constant	-3.967*** (0.807)	-1.147*** (0.217)	0.408* (0.168)
Adjt. R-squared	0.528	-	-
Pseudo R-squared	-	0.178	0.273
Log Likelihood	-	-2488.048	-731.427
# Obs		1,452	

*** 0.001; ** 0.01; * 0.5; † .01;

Things I'm Currently Working on:

- Use more schools (Moody's [2001] *Countryside High* and *Mountain Middle school* as extreme examples of variation in friendship segregation along racial lines).
- Include noise (Macy and Tsvetkova 2013). Randomly distributed thresholds & the ability to make mistakes
- Include more boundaries (DiMaggio and Garip 2011)
- Explore differences between strongly vs weakly symmetrized nets

In Summary

- I propose to examine diffusion processes driven by liminal individuals positioned in-between bright boundaries.

- Liminal individuals (e.g. multiracial individuals) could ameliorate the effects of ethnoracial segregation in adolescent networks in the context of the diffusion of innovations.

Thank you!

Appendices

The Nature of the Data

Theoretically, Hartup and Stevens (1997, 1999) argue that there is a deep structure and a surface structure to friendship. They refer to the deep structure of friendship as the *essence* of the relationship, while considering the surface structure dependent on the particular social exchanges that are typical of a specific developmental stage. According to Hartup and Stevens, the essence (i.e. deep structure) of friendship in Western cultures is *symmetrical reciprocity* (see also Bukowski and Hoza 1989).

The Nature of the Data

A wealth of evidence based on friendship nomination data has shown the clear existence of a positive reciprocity effect in this kind of data, that is, adolescents do exhibit a general tendency to reciprocate an incoming friendship nomination (e.g. Leszczensky and Pink 2015; Snijders and Baerveldt 2003; Gesell, Tesdahl and Ruchman 2012; Moody 2001).

This supports the the basic –albeit fundamental– idea that friendship is a relationship in which mutual acknowledgement is key (Kitts 2014).

Computational experiments and observational evidence show that reciprocity may be the key component to distinguish human nets from technological nets (Schnegg 2006).

The Nature of the Data

I am not the first one to assume reciprocity as a necessary condition to analyze friendship nomination data (Parker and Asher 1993; Wentzel et al. 2004, Reiter-Purtill et al. 2010; Berndt and Perry 1986; Bagwell et al. 1998; Goodreau et al 2009; Hruschka 2009).

However, most scholars analyze friendship nomination data *without* using reciprocal nominations as a way to identify friendships (Moody 2001; Hallinan and Williams 1987, 1989; Hallinan and Kubitschek 1990a, 1990b; Quillian and Campbell 2003; González et al. 2007; Kao and Joyner 2004, 2006; Smith et al. 2016; Clark and Ayers 1992).

The Nature of the Data

- Data: Add Health,
  Friendship nomination data
- Ethnoracial homophily
  (Kao and Joyner 2006; Hallinan 1978; Moody 2001; Goodreau et al. 2009)
- Missing? Aspirational? New or unstable?
  (Ball and Newman 2013)

Weak symmetry: if A → B or A ← B, then A ⟷ B

Adoption by Key Boundaries (T=0.25,R=10,p=0.1)

Features of seed agent	Average Adoption In-group Alters (A)	Average Adoption Out-group alters (B)	Difference (A – B)
Race	0.452	0.346	0.106
Race by Interviewer	0.454	0.343	0.112
Ethnicity*	0.396	0.347	0.050
Grade	0.376	0.362	0.014
SES	0.368	0.368	0.000
Sex	0.387	0.366	0.001
Age	0.376	0.365	0.012
Overall Average Adoption		0.368	

*Latino= 1; otherwise = 0

Average Adoption within the Racial boundary (T=0.25,R=10,p=0.1)

Race of seed agent	Average Adoption In-group Alters (A)	Average Adoption Out-group alters (B)	Difference (A – B)
Black	0.538	0.192	0.346
White	0.405	0.300	0.105
Native American	0.557	0.447	0.110
Asian	0.488	0.452	0.036
Multiracial	0.396	0.341	0.054
Other	0.459	0.348	0.111
Latino (any race)	0.443	0.310	0.133
Overall Average Adoption		0.368	

*Latino= 1; otherwise = 0

So What? Ethnoracial Liminality & Health

- Acculturation

  (David and Katzman 1999; Altman et al. 2017; Gorman et al. 2016)

- The literature on (adolescent) friendship networks and health has not fully engaged in a deeper understanding on the interaction between networks, race & ethnicity, and health.

  (Salvy et al 2012; Mcdonald 2012; Fletcher 2011; Karp and Gesell 2015).

So What? Ethnoracial Liminality & Health

- Are culturally liminal individuals associated with wide patterns of diffusion of health-related behs?

- Obesity, screen time, playing sports.

- Stochastic Actor-Based Model (Snijders 1996).

Show that liminal agents are important for diffusion processes above and beyond well-known homophily and selection effects

Homophily-Based ERGM for Sunshine (Log-odds)

Term	Estimate	Std. Error	p-value
Edges (Intercept)	-10.894	0.098	***
Edgewise shared part.	0.892	0.033	***
Sqrt(degree)	0.782	0.019	***
Same gender(men)	0.459	0.042	***
Same gender(women)	0.440	0.042	***
Same grade(10th grade)	1.287	0.050	***
Same grade(11th grade)	1.071	0.047	***
Same grade(12th grade)	1.333	0.054	***
Same age(14 yrs)	0.291	0.091	**
Same age(15 yrs)	0.151	0.058	*
Same age(16 yrs)	0.119	0.053	*
Same age(17 yrs)	0.326	0.078	***
Same age(18 yrs)	1.152	0.302	***
Same race(white)	1.290	0.079	***
Same race(black)	2.121	0.058	***
Same race(asian)	1.276	0.041	***
Same race(other)	1.307	0.061	***
Same race(multiracial)	1.098	0.259	***
Same SES(low)	0.171	0.063	**
Same SES(medium)	0.081	0.045	†
Same SES(high)	0.166	0.068	*
Dyads = 2,459; Nodes = 1,452; BIC 34912			

'Native American' dyads; 9th grade dyads; and19 years-old dyads are excluded due to small sample sizes.
*** 0.001; ** 0.01; * 0.5; † .01;

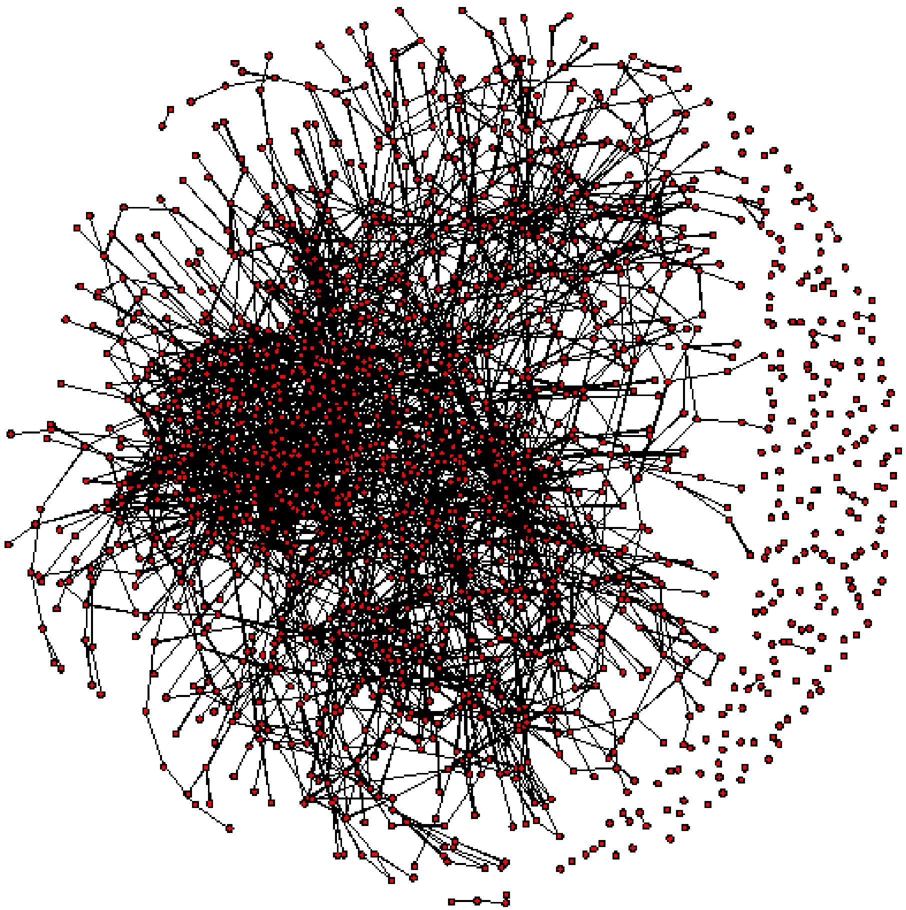
Full Network (Weakly Symmetrized)

Nodes = 1676

Ties = 3565

Isolates = 193

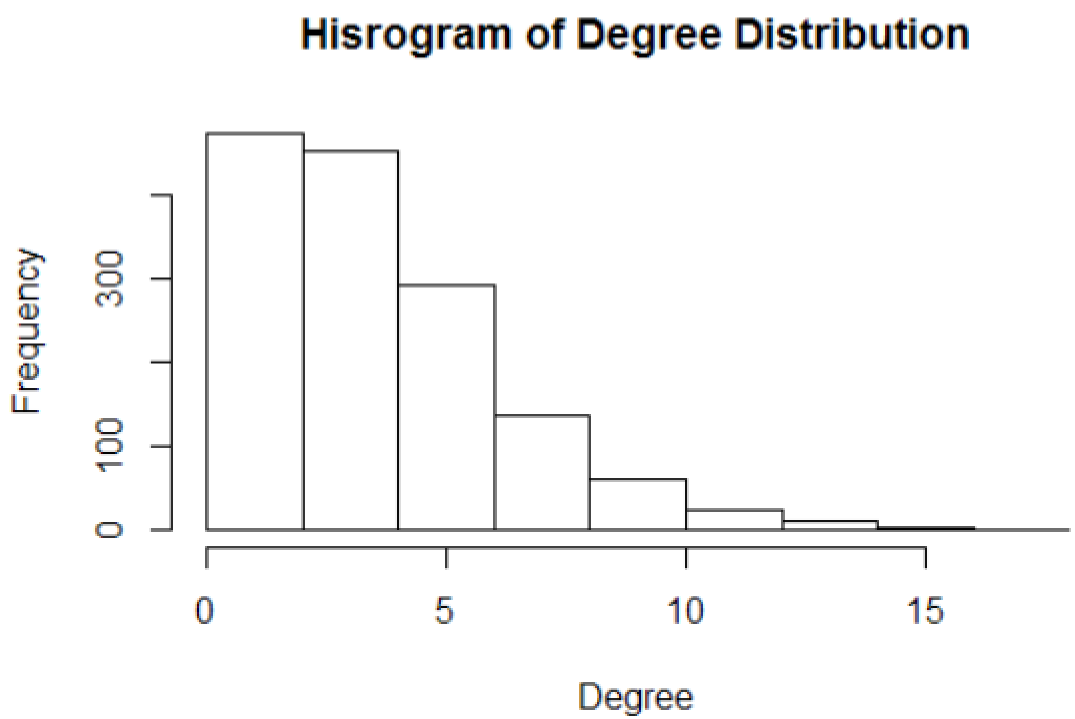
Components of size > 1:
Size 2 = 9
Size 3 = 3
Size 4 = 1



Distribution of Cliques by size

Clique Size	Frequency
1	0
2	1483
3	452
4	106
5	5
6	1

Degree Distribution



Quantiles:

- 0% =1
- 25% = 2
- 50% = 4 (median)
- 75% = 5
- 100% =18

Mode = 2

Mean =4.1

SD = 2.6

degree $\leq 2 = 32.4$ %

degree $\leq 3 = 49.5$ %

degree $\leq 4 = 63.8$ %

A Critical View on Liminality

1. Do not romanticize liminality
   (Anzaldúa 1987; Huang et al 2008)

2. Do not naturalize liminality
   (Ang 2001; Emirbayer Desmond 2015)

3. Do not universalize liminality
   (Davis 1991; Bonilla-Silva 2002)

Liminality, and multiracialism more specifically, should not be considered as *the* solution to the entrenched inequalities on which the US racial formation rests. Part of the reason why this is true is because the idea of multiracialism is not a negation of race, it is, on the contrary, rooted on the idea of race (Telles and Sue 2009; Emirbayer and Desmond 2015; Dacosta 2007).