Modeling Sociocultural Influences on Decision Making: Understanding Conflict, Enabling Stability,
Joseph V. Cohn, Sae Schatz, Hannah Freeman, David J. Y. Combs, eds,
CRC Press, Taylor and Francis, pp. 61-84, 2016.

CHAPTER 4

THE SPREAD OF OPINIONS IN SOCIETIES

Boleslaw K. Szymanski, Omar Lizardo, Casey Doyle, Panagiotis Karampourniotis, Pramesh Singh, Gyorgy Korniss, Jonathan Bakdash

Abstract

We demonstrate the role of sociocultural factors in understanding opinion dynamics. We start by assessing theoretical models of opinion spread whose parameters are dependent on sociocultural factors. First, we model the dynamics of opinion spread in communities with both static social and dynamic social networks using the Binary Agreement (BAM) and Threshold (TM) Models to recreate the evolution empirically observed in the data. Then, we examine the unique set of behavioral network data on social interactions (based on electronic logs of dyadic contact via smartphones) collected at the University of Notre Dame. The participants are a sample of members of the freshmen class of 2011 whose opinions on political and social issues have been regularly recorded—at the beginning and end of each semester—for the last three years. Then, we measure the evolution of participants' opinions and ascertain the quantitative dependence on the cultural traits of individuals and the structural properties of social networks that they form. Our analysis shows that ties among people who are more likely to share opinions (e.g. of the same race, gender, or socioeconomic class) decay slower than ties among persons who are likely to have different opinions. The analysis also indicates that the partner selection of individuals is correlated with sharing a (political) opinion. Results show sociocultural factors and social network dynamics both influence the evolution of opinions in multicultural social networks.

1. Introduction

Here, we focus on the two most popular models of opinion in social networks, the Threshold Model (TM) (Granovetter 1978; Watts 2002) and the Naming Game Model (NG) (Steels 1995; Baronchelli et al. 2006; Baronchelli et al. 2008; Dall'Asta et al. 2006). In TM, an individual adopts a new opinion only when the fraction of its nearest neighbors possessing that opinion is larger than an assigned threshold, which represents the resistance of the node to peer pressure. In the naming game (NG) model, each individual in the network possesses a list of opinions which comprise its opinion state. In the version of the NG model considered here each interaction starts with randomly selecting the speaker first, and then randomly choosing the listener among the speaker's neighbors. The speaker then shares a randomly chosen opinion from its opinion state with the listener.

2. Naming Game Based Models

One of the simplest and most popular models describing large scale patterns is the Naming Game (NG) model (Steels 1995). The rules governing these local interactions are as follows. Every node (or individual) in the system contains a list of words. At the beginning of each time step, a speaker and a listener connected to it are randomly chosen from these nodes, along with one word from the speaker's list. If the listener's list contains that word, both the speaker and the listener will delete all other words from their lists. If the listener does not already have that word in its list, it will simply add it with no deletion of words on either side.

Here, we use an even simpler model termed the binary agreement model, in which only two words *A* and *B* are present and nodes transition between three possible states, *A*, *B*, and *AB* (Castelló et al. 2009, Xie et al. 2011). The reduction in the number of possible states leads to a drastically reduced time to consensus, now on the order of $ln(N)$ (Baronchelli et al. 2008, Castelló et al. 2009, Zhang et al 2011). Here, by construction, the states are entirely symmetric with no bias towards either opinion. The game changes with an addition of committed agents who are set to a singular opinion (for simplicity assumed to be *A*) that will never change (Lu et al. 2009, Xie et al. 2011, Zhang et al. 2011). They are still able to influence the opinions of the other nodes around them. In the simplest and most direct case, the system starts with some predetermined number of committed agents *p*, while the rest of the population are all in the opposite opinion *B*. For a finite network given infinite time, a system such as this will always reach consensus on the opinion favored by the committed agents. If the initial population of committed agents is small, a large random deviation is required to achieve consensus, so the time to consensus for such systems is in the order of e^N. However, there is some critical population fraction denoted p_c that causes a sharp phase transition in the system (Xie et al. 2011). Above p_c, the time to consensus is quite fast, on the order of $ln(N)$. A critical population in a fully connected graph is $p_c=0.09789$ (Xie, et al. 2011). Above this population there exists only one absorbing state at the consensus for *A*. Below, however, there is the consensus absorbing state in addition to an active steady state and a saddle point.

The model can be extended to two competing committed groups: one in state *A* and one in state *B*. Here the populations of the committed agents will be termed p_A and p_B respectively (Xie et al. 2012). The system can enter into different phases depending on the initial setup of the critical population sizes. With these populations initially equal, the system has two active steady states and one saddle point. When the system is in either state, it will take a large fluctuation of opinions of non-committed nodes to push the minority population past the saddle point and into the second steady state. When these populations are unequal and below critical values, the saddle point will move to favor the population with the higher number of committed agents. In these cases, the average time to switch from one state to the other is on the order of e^N since a large random fluctuation is needed for such switch. Finally, when the populations fall outside of their critical values then the system maintains only a single stable state where the population with the higher number of committed agents dominates. In these cases, the system consensus is reached quickly on the time scale of $ln(N)$.

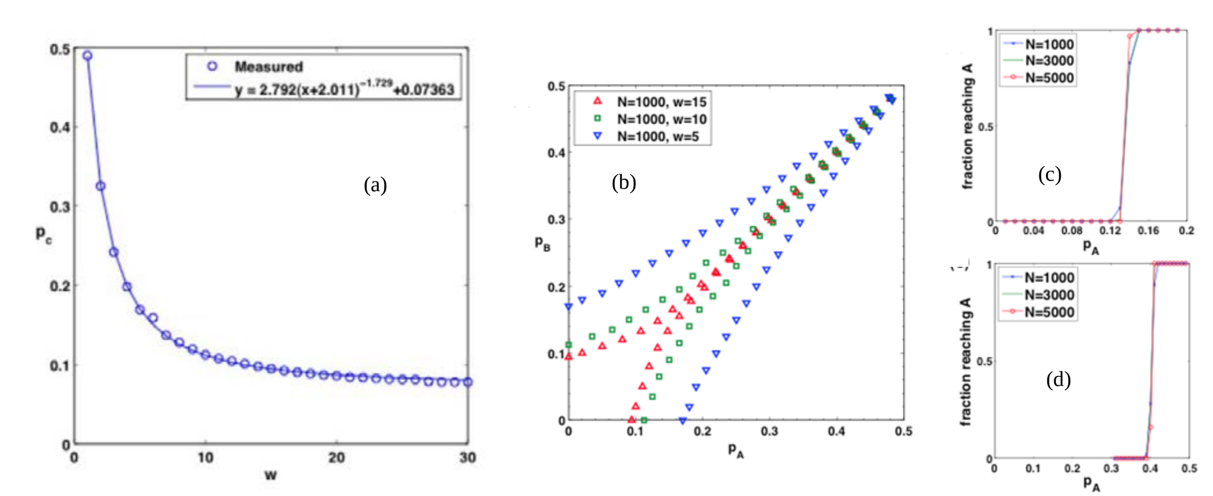

Figure 4.1. (a) The critical committed fraction $p_c(w)$, as a function of commitment strength w for Barabási-Albert networks of size $N = 1000$ with average degree $<k> = 10$ (Galehouse et al. 2014). (b) Phase diagram of the system in the presence of competing committed groups holding opinion A and B with initial sizes p_A and p_B, respectively for Barabási-Albert networks with average degree $<k> = 10$. (c) The fraction of runs reaching the all-A consensus state along a trajectory $p_A + p_B = 0.2$ in the parameter space for $w=10$ and $<k> = 10$. (d) The fraction of runs reaching the all-A consensus state along a trajectory $p_A + p_B = 0.8$ in the parameter space for $w=10$ and $<k> = 10$ (Galehouse et al. 2014).

Next we consider *waning commitment* so nodes holding it move into the mixed state after w consecutive interactions with a node of the opposite opinion. With the ability of the committed nodes to now become normal, the system regains the possible absorbing states for A and B. With a single group of weakly- (or partially-)committed agents present, the critical population of weakly-committed agents required to force a consensus varies with w, the extent of which is seen in Fig. 4.1(a) (Galehause et al. 2014) and can be numerically approximated with $\alpha \cong 1.73$ by

(4.1) $$p_c(w) \cong p_c(\infty) + \frac{a}{(b+w)^{\alpha}}.$$

Next, we consider competing weakly-committed agents. As seen in Fig. 4.1(b), 'beak' emerges where the system will undergo a phase transition when p_A and p_B cross its threshold. The value of w will change the width of the beak, with higher w imposing narrower beak. In addition, the stable points previously present inside the beak are now absorbing states since the committed population can eventually be swayed.

The idea of waning commitment introduces stubbornness into the system, where the nodes can resist the change without becoming implacable as is the case with the fully committed nodes. An extension of this behavior is to make it not inherent to a set of nodes but a part of the opinion enabling it to spread with the opinion. This extension requires eliminating an intermediate state AB (Doyle et al. 2015). Instead, each state is assigned an 'idea inertia' denoted w_A and w_B, respectively, corresponding to how many consecutive times a node in that state must hear the opposite opinion before it switches. If a node listens to an opinion that confirms its own, its count towards switching is reset to zero. Alternatively, if it hears and opinion that is contrary to its own, it increases its count. Once the count reaches node's inertia, it switches states and reset its count at zero. When the system is set up similarly to the committed agent systems (with a minority population of opinion *A* with varying w_A and a majority of opinion *B* with constant w_B), the critical population of opinion *A* initiators behaves similarly to that in the waning commitment model. This can be seen in Figures 4.2(a-b) (blue circles), where $w_B=2$ for every simulation. Calculating rate equations for waning commitment models is difficult because they are non-Markovian. Fig. 4.2(b) shows the dynamics, where the numerical solutions capture the qualitative behavior of the system quite accurately, yet consistently overestimates the value of p_c. Nevertheless, such an upper bound is useful as it is a value at which the system will always reach consensus for the minority state given that inertial value.

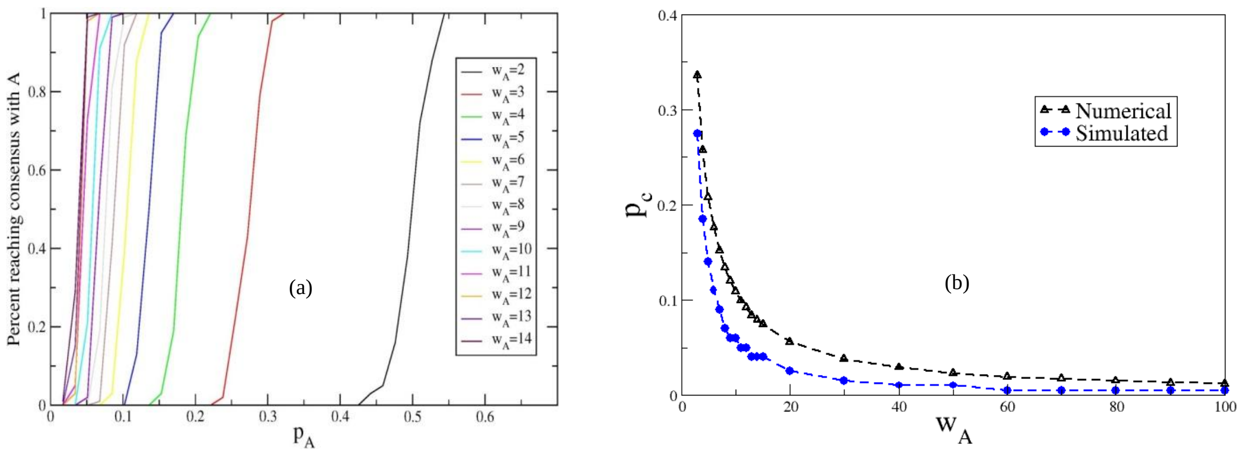

Figure 4.2. (a) The fraction of simulation runs that reach consensus on opinion *A* vs. the initial density of agents with opinion *A* on a complete graph with $N=1000$ and $w_B=2$ (Doyle et al. 2015). (b) . Individual-based simulation results (blue circles) and numerical solutions (black squares) for the critical fraction of committed minority p_c as a function of waning commitment w_A.

3. Threshold Models

One of the simplest models that capture the adoption dynamics of a new opinion, behavior, trait, product etc. is the threshold model (TM) (Granovetter 1978; Centola et al. 2007; Singh et al. 2013). In this model, a node's opinion is represented by a binary variable ('active' or 'inactive'). An inactive node becomes active (i.e., it adopts the new opinion) if at least a threshold fraction (φ) of its neighbors are already active. The system continues to evolve till a steady-state is reached in which no activations are possible. The adoption threshold φ can be the same for every node, or it can be drawn randomly from a distribution. The dynamics is also asymmetric since an active node stays active indefinitely. The relative size of the active population in the eventual steady-state is also called the cascade size (S_{eq}) (Watts 2002; Singh et al. 2013).

3.1. Tipping Points in Threshold Models

First, let's assume that all nodes have an identical threshold φ. A small set of nodes holding a new opinion (initiators) can cause activation cascades and even convert the population to the new opinion (global cascades). The structure of the underlying social network plays an important role in determining whether or not such cascades can be triggered by a single initiator (or small single-clique initiators).

For Erdős-Rényi (ER) networks, global cascades cannot be triggered if the average degree of the network is too low (because then the network consists of small isolated clusters) or too high (because it is difficult for the high degree nodes to satisfy the threshold condition). The critical threshold φ_c, above which cascades cannot be triggered, is inversely related to the average degree $<k>$ (Watts 2003). Therefore, global cascades can arise only in the intermediate range (cascade window) of $<k>$. Fig. 4.3(a) shows the simulation results of the average cascade size (S_{eq}) as a function of average degree $<k>$ for a fixed initiator fraction (initiators are selected at random).

In real social systems adoption thresholds can be very high (Latane and L'Herrou 1996) since individuals are only likely to change their opinion if at least half of their interacting neighbors hold the opposite opinion.

A careful selection (e.g. influence maximization algorithms, heuristic strategies (Singh at al. 2013, Morone and Makse 2015)) of the initiators can trigger global cascades, when randomly selected initiators cannot, but they are ineffective when φ is very high. However, global cascades can be triggered by increasing the initiator fraction p once it reaches a 'critical value' p_c. At that point, a discontinuous transition occurs and large cascades are seen immediately, as shown in Fig. 4.3(b). The relevant quantity, $\tilde{S}_{eq}$, discounts initially active nodes and represents the fraction of nodes initially inactive that eventually adopted active state. Transitions in $\tilde{S}_{eq}$ are shown in

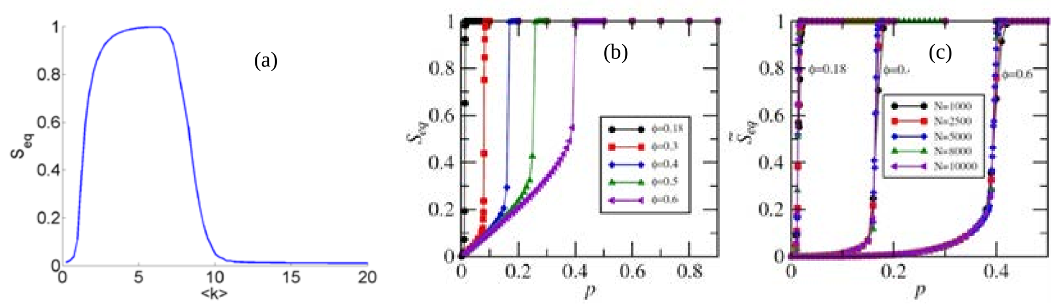

Figure 4.3. (a) The cascade window for ER networks, *N=1000*, *φ*=0.18 and initiator fraction *p=0.01*. Cascade size and scaled cascade size as a function of initiators on ER networks with *<k>=10*. (b) Cascade size S_{eq} as a function of initiators *p* for ER networks with *N=10000* for different values of *φ*. (c) Scaled cascade size $\tilde{S}_{eq}$ vs. *p* for ER networks with different network sizes *N* and *φ* values (Singh et al. 2013).

Fig. 4.3(c) for different φ values and several network sizes; clearly the transition depends only on φ and is independent of system size N. This transition (the emergence of the tipping point) is quite generic in the threshold model, and can be observed in networks with different sizes and average degrees, as well as for different selection methods for initiators (Singh at el. 2013).

3.2. Dynamics with Individualized Thresholds

The empirical studies (Latane and L'Herrou 1996, Centola 2010) showed that the threshold varies from node to node necessitating different thresholds at different nodes to capture the complex nature of social influencing.

A natural choice for the distribution of the adoption thresholds is the normal distribution, but then a node has non-zero probability having threshold higher than one, making it too stubborn to adopt a new opinion, or negative threshold, making it self-activating instigator (Granovetter 1978). An interesting case arises when instigators but not stubborn nodes are present; with large enough threshold distribution, a cascade can occur without any initiators, mimicking spontaneous riots started by instigators for which an accurate analytical approximation was introduced (Gleeson and Cahalane 2007).

Excluding the above case, we consider cascades that are initiated only with initiators (Karampourniotis at el. 2015). We assume that every node can potentially be an initiator, and that thresholds are limited to range (0,1). We consider the impact of the diversity of the threshold drawn from a truncated normal distribution on the cascade size in the presence of multiple initiators. The mean threshold φ_0 is kept constant at 0.50, while the standard deviation σ varies from *0*, to *0.288*. For the simulations, we randomly assign initiators one by one and measure the cascade size. We repeat this process by drawing thresholds from the same distribution. The cascade size S_{eq} for each threshold distribution is obtained by averaging one thousand times on different threshold distribution draws.

Figures 4.4(a-b) show the effect of σ on the cascade size S_{eq} for a constant initiator fraction p and constant mean threshold φ_0. As σ increases so does a fraction of nodes whose threshold is far from the average, causing a twofold effect. Those with thresholds far below average are easily activated while those with thresholds far above average are increasingly difficult to activate. Thus, when the initiator fraction is small, S_{eq} is monotonically increasing following the increase of fraction of low threshold nodes. However, when the initiator fractions are large, this increase helps a little since those nodes are likely to be already activated which leads to the non-monotonic behavior seen in Figures 4.4(a-b).

Following (Karampourniotis at al. 2015), we study how the cascade size S_{eq} varies with the initiator fraction p. Figures 4.4(c-d) show that increasing the initiator fraction, for small enough σ, transitions small local cascades to large global cascades which, for synthetic networks is a discontinuous phase transition; however, averaging over different network ensembles smoothens the transition point (see Fig. 4.4(c)). For larger σ the initiator fraction for which the transition occurs is reduced, while for the synthetic networks the spread size still exhibits a discontinuous phase transition. With largely diverse thresholds, a critical initiator size, beyond which cascades become global, ceases to exist and the tipping-point behavior of the social influencing process disappears. It is replaced by a smooth crossover governed by the size of initiators. This property can be important for example, for a company's marketing strategy of a new product. If the threshold distribution is narrow enough, unless a critical initiator fraction is reached, there is a marginal local spread on a few of the first or second neighbors of the initiators. On the other hand, if the threshold distribution is wide, there is a significant spread.

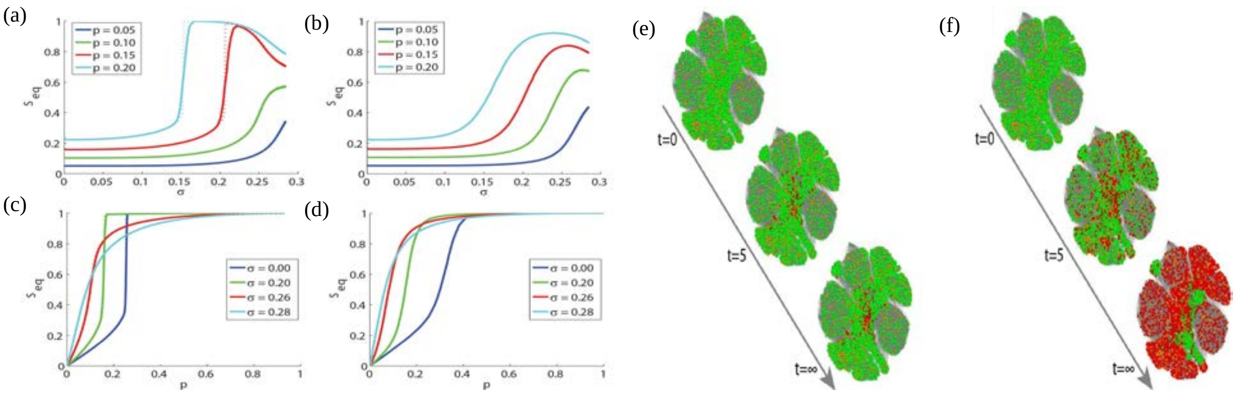

Figure 4.4. Behavior of the cascade size at equilibrium S_{eq} for ER networks for varying standard deviation σ for (a) ER networks and (b) the FB network, and for varying initiator fractions p for (c) ER networks and (d) the FB network. The ER networks structure is $z=10$ and $N=10000$. The FB network structure is $z=43$ and $N=4039$. The mean threshold is $\varphi_0=0.50$. The simulations (solid lines) are averaged over one thousand repetitions, while the analytic estimates (dotted lines) are based on the tree-like approximation (Gleeson and Cahalane 2007)). (e,f) Visualization of the spread of opinion in the TM model on the FB ego-network with the fraction of the randomly selected initiators at $p=0.20$. The standard deviation of the threshold is (e) $\sigma=0$, (f) $\sigma=0.20$. Inactive nodes, initiators, and active nodes (through spreading) are marked with green, orange, and red, respectively (Karampourniotis et al. 2015).

For ER networks the analytic estimates based on the treelike approximation model (Gleeson and Cahalane 2007) are in good agreement with the simulations. To show this behavior for real networks, we use the TM for a Facebook ego-network (FB network) (SNAP) in Figures 4.4(b, d). In this case, the cascade size change is smoother than for the ER networks because communities in the FB network have different critical fractions of initiators. A visualization of the spread evolution in the FB network (see Fig. 4.4(f)), shows the existence of three communities which are not activated through spread. Finally, a comparison of Fig. 4.3(e), where $\sigma=0$, and Fig. 4.3(f), where $\sigma=0.20$, clearly shows that the spread is greater for larger σ (Karampourniotis at el. 2015).

4. Impact of Cultural and Gender on Spread of Opinion in Empirical Networks

4.1 Notre Dame NetSense Data

Mega-city Xanadu is a very large-scale social system whose underlying interactions are structured by cultural, racial, and class divisions. As such, this example, while fictional, resembles the way that most real life social systems are organized. As has been well-known for decades in sociology and anthropology, persons organize their interactions along any distinguishable physical, cultural, biological, and socially constructed characteristic, with those related to age, gender and ethnicity (especially when combined to class) having the most influence on how interacts with whom. For instance, Xanadu, but first being open to migration and then closing its border created a correlation between place of origin and class location for North Razini immigrants. These immigrants are more likely to live in segregated areas, are more likely to interact with co-ethnics and are less likely to form connections with members of the majority. These are dynamics that unfortunately mirror those of ethnic minorities struggling to assimilate throughout the developed and developing world. As such studying the link between interaction and exogenous socio-cultural markers (such as gender and race) becomes important for understanding these large-scale patterns. This matters even more if it is likely that the link between social characteristics and dynamics of social interaction affect the way in which we understand opinion convergence in large-scale social systems and the persistence of minority subcultures unlikely to be swayed by the majority culture.

We set out to test an overall theory linking "homophily" (preference for similar others) and the dynamics of tie formation and decay in real social systems. The study real on a small scale but ecologically natural setting: Adolescents making the transition to college in a residential campus. Data are obtained from the ongoing *NetSense* study at the University of Notre Dame (Striegel et al. 2010; Wang, Hachen, and Lizardo 2013). The study equipped a cohort of roughly 200 incoming first-year students with smartphones and tracked, among other things, the calls and texts made and received (but not the content of their communications) over a two-year period using a monitoring app employed on each phone. This app logged and then transmitted to a secure database a Call Data Record (CDR) for each communication event. Each CDR contains the phone numbers of the sender and receiver along with a timestamp indicating when the event occurred. While we have thousands of CDRs for texts and calls involving a NetSense subject and people outside the study, we use only the CDRs in which both parties are in the NetSense study, because for both parties involved in the communication we have data derived from surveys on their socio-demographic characteristics, tastes, and opinions as well as other variables of interest derived from a survey administered to all NetSense subjects prior to their arrival on campus. However, the activity analysis is based on the entire set of outgoing communications for each participant (complete ego-network degree) whether the recipient was within the study or not.

4.2 Impact of Agent Socio-Demographic Characteristics on Activity and Assortative Mixing

In this section we first summarize the findings concerning socio-demographic differences in both activity and sociability and assortative mixing (Newman 2003). Sociability effects refer to the higher likelihood of members of certain categories to form more connections to other persons regardless of the socio-demographic traits of those persons. Studying assortative mixing in the context of modeling opinion spread is important for the following reasons. It is well know that opinions, beliefs, and other cultural commitments tend to be different for members of different socio-demographic groups. This means that when persons interact with those who are similar to them, their current beliefs and values are likely to be reinforced. For instance, in Mega-City Xanadu, the North Razinis “struggle to mesh” with the society seemingly both because they are rejected by others but also because they have higher levels of ingroup homophily. This makes them different from the Bursuks who are weaker homophily patterns and are thus more likely to consider interacting with outsiders. Opinion change in real social networks is more likely to happen for those individuals who are more likely to venture outside of their own group for sociable interactions (Aral et al. 2009). A social system characterized by high levels of assortative mixing is also one that is likely to exhibit opinion polarization rather than homogeneity.

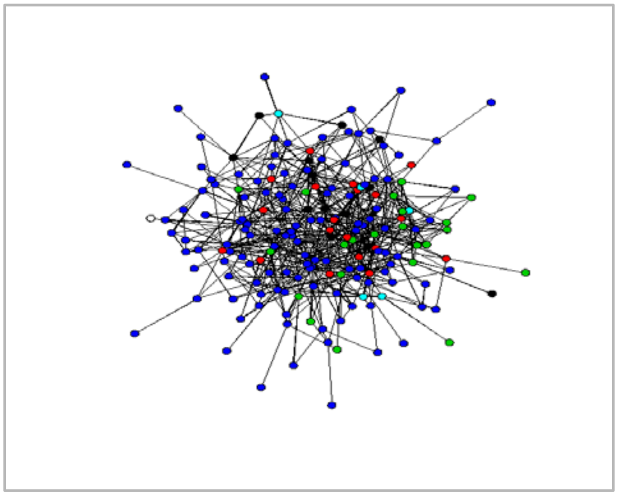

Figure 4.5. NetSense within-study network. Vertex color indicates race, Blue = White, Red = African American, Green = Latino, Black = Asian, Teal = Other.

We identify both sociability and assortative mixing effects using exponential random graph models (ERGMs) fitted to the network (Robins et al. 2007). These models allow us to ascertain statistical dependencies on the network while adjusting for lower order network properties, such as the overall tendency to form edges (density) and the tendency for agents who share neighbors to also be connected to one another. The purpose of ERGMs is to provide a summary description of the local relationship formation mechanisms that generate the global structure of the network. For instance, the *NetSense* within-study network (depicted in Fig. 4.5), may be considered similar to the response or outcome variable in a regression or classification model. Here the predictors of whether we observed an edge (or not) are such things as "the propensity of individuals of the same gender/race/religion/class background to form connections" or "the propensity of individuals who share friends to be connected to one another." In Fig. 4.5 for instance, it is clear that individuals who share the same racial/ethnic status tend to cluster together. The ERGM can help us quantify the strength of this assortative mixing effect, while holding constant other confounding relationship formation mechanisms.

The exponential random graph model specifies the conditional log-odds of an edge existing between two nodes vs. not existing, conditional on values of covariates and a network configuration:

(4.2) $$\frac{\Pr(X_{ij}=1|Y=y,X_{ij}=x_{ij})}{\Pr(X_{ij}=0|Y=y,X_{ij}=x_{ij})} = e^{\theta^T\delta_{ij}(x)+\theta_a\left(y_i+y_j\right)+\theta_b y_i y_j}$$

where θ_a denotes the (additive) activity parameter and θ_b is the (multiplicative) assortative mixing parameter; y_i, y_j are the values of each vertex on a given socio-demographic trait of interest (in this example gender, which is a categorical trait with two values: male, female). The other term is for the configuration (endogenous) part of the model. In our case, this involves a parameter for the overall number of edges in the graph (density) and a parameter that counts the number of shared neighbors for each node in the graph (transitivity). For the purpose of interpreting the activity and mixing parameter, we can ignore that part of the model because when taking odds ratios it drops out (assuming that the configuration model is the same for all edges).

What we can compute from this model is the conditional odds ratio. The reference category for these odds is $y_i = y_j = 0$, i.e. the (0,0) cell in the $n \times m$ table cross classifying edges by whether they have or do not have the trait $Y = y$, which in our case is a 2 x 2 table cross-classifying edges by the gender mix of the vertices attached to them. Thus, for the example used here Y is gender and $Y = 1$ indicates that the node is female.

The conditional odds ratio for an edge having one node with $Y = 1$ versus having both vertices with $Y = 0$ (e.g., odds ratio of male-female tie versus male-male tie) is e^a, the activity parameter. Its value defines the odds that, conditional on one node being male, the other will be female. If the activity parameter is 0, then the odds ratio is 1, making the odds of male-female tie the same as the odds of male-male tie. This means the counts of male-male and male-female ties should be the same. If the activity parameter is greater than 0, then male-female ties are more prevalent than male-male ties, if less than 0, male-male ties are more prevalent than male-female ties. The activity parameter reflects difference in the degree of male versus female, because the only way male-female ties can be more prevalent than male-male ties (assuming random matching, i.e. no homophily) is if females have more ties than males.

If there is no homophily (i.e., random matching), then the odds of a female-female tie versus a male-male tie is e^{2a}. That is we expect more female-female ties solely because there are females having more edges (i.e. are more likely to be at the end of a random edge), so that the odds of a female being at both ends of an edge is the exponential of 2 times the activity parameter. If there is homophily (positive b parameter) then the odds of a female-female tie versus a male-male tie is e^{2a+b}. The $2a$ captures the greater odds of a female-female tie due to the higher activity of females, while the $_b$ term adds in increased odds of female-female ties resulting from assortative mixing. Note that the odds of the female-female tie versus a male-female tie is now e^{a+b}. Relative to the male-female edge, the female-female edge gets a boost over the male-female edge because females have higher activity and because of activity differences female-female ties are more prevalent.

For instance, from our estimates (Table 4.1, Model 4) we can see that a= 0.14 and b= 0.46. We coded female as 1 so the reference like above is male-male ties. The odds of a male-female versus a male-male ties is $e^{0.14}$ - 1 = 15%, i.e., the former are 15% more likely than the later, reflecting the higher activity of women. Without the homophily effect this would imply that female-female ties would be about 32% more prevalent than male-male ties (e.g. $e^{2*0.14}$ - 1). But taking into account the assortative mixing effect and the odds of a female-female vs. a male-male tie is $e^{2*(0.14 + 0.46)}$ = 2.09, implying that female-female ties are more than twice as likely 209% more likely) than male-male ties. Finally, contrasting female-female versus male-female ties, we see that female-female ties are $e^{(0.14 + .46)}$ = 1.82 times more likely that male-female ties.

In our estimates (Table 4.1) female-female ties are more prevalent, than male-female ties, which are more prevalent then male-male ties. This is indicative of strong female same-gender preference effect coupled with females having high degree, which implies a tendency for males to not associate with other males because they randomly associate with other people who happen to be more likely to be females because of their higher activity.

Table 4.1: Gender-Based Activity and Assortative Mixing

	Model 1	Model 2	Model 3	Model 4
Density	-3.24	-3.76	-4.15	-4.51
	(0.04)	(0.08)	(0.08)	(0.10)
Activity (Women)		0.25		0.14
		(0.05)		(0.04)
Mixing		0.51		0.46
		(0.08)		(0.07)
Transitivity			0.99	0.97
			(0.07)	(0.07)

Race and Ethnicity - Table 4.2 shows the result of fitting and ERGM model with separate activity parameters for each racial and ethnic group and a common assortative mixing parameter for all race groups. We find that, in terms of activity, minority groups are more active than Whites in their first semester on campus. Recall that the activity coefficients compare the odds of observing an edge featuring a non-white respondent versus the odds of observing an edge containing two white respondents. Across all racial groups (as indicated by the positive activity effects) we find that mixed race edges are more prevalent than the all-White edges.

Table 4.2: Race-Based Activity and Assortative Socializing

	Model 1	Model 2	Model 3	Model 4
Density	-3.24	-4.53	-4.15	-5.13
	(0.04)	(0.10)	(0.08)	(0.10)
Activity (Blacks)		0.99		0.75

		(0.10)		(0.08)
Activity (Latinos)		0.84		0.64
		(0.08)		(0.07)
Activity (Asians)		0.79		0.59
		(0.08)		(0.06)
Activity (Other)		0.75		0.64
		(0.19)		(0.13)
Mixing (All Races)		1.33		1.12
		(0.10)		(0.08)
Transitivity			0.99	0.95
			(0.07)	(0.07)

Class Background.- We find that (Table 4.3, Model 3) there are no activity differences based on class background, once we adjust for the statistical propensity to find class-matched edges over class mismatched ones. In addition, we find that same race dyads are statistically more prevalent than mixed-race dyads (= 1.33), even after adjusting for the propensity of friends of friends to also be connected to one another (= 0.95). Thus, even though Whites are less active than non-whites, when they do form ties, they are disproportionately likely to be directed at other Whites. Non-whites on the other hand, have similar same-race preferences, but due to their higher activity, end up forming more mixed-race connections than we would expect by chance. In our data, we coded class background as follows: first generation college students were coded as working class ($_{b}$ = 1.06). Neither middle class nor upper class students display any tendency to

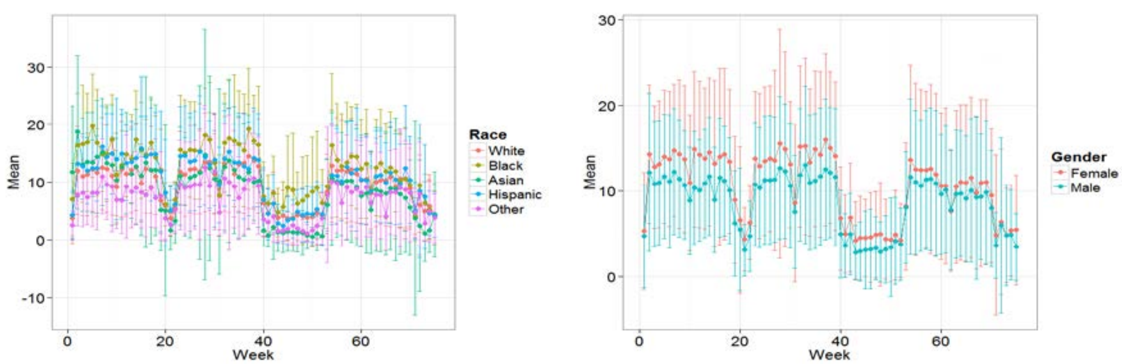

Figure 4.6. Race and gender effects on active network size.

form ties disproportionately with members of the same class group. Note that the working class affinity effects persist even after adjusting for the tendency towards transitivity.

Table 4.3: Class-Based Activity and Assortative Socializing

	Model 1	Model 2	Model 3	Model 4	Model 5
Density	-3.24	-3.71	-3.56	-4.15	-4.17
	(0.04)	(0.14)	(0.21)	(0.08)	(0.08)
Activity (Working Class)		0.50	0.21		
		(0.10)	(0.12)		
Activity (Upper Middle Class)		0.11	0.13		
		(0.09)	(0.19)		
Mixing (All Classes)		0.31			
		(0.10)			
Mixing (Working Class)			1.06		0.90
			(0.27)		(0.14)
Mixing (Middle Class)			-0.44		
			(0.38)		
Mixing (Upper Middle Class)			0.11		
			(0.22)		
Transitivity				0.99	0.98
				(0.07)	(0.07)

Also of interest is time evolution of network activities for different classes of agents, classified according to their personality, gender and race. The resulting diagrams, shown in Fig. 4.6, provide also range of differences between the respective classes of agents as well as trends over time. The strong personality, race, and gender effects on active network size of students after joining the university evolve differently. As shown in Fig. 4.6, race effects persist through the observation period but gender effects decline.

4.3 Network/Opinion Co-Evolution: Differential Tie Decay by Dyadic Mix

A key implication of the dynamic naming game model is that the opinion evolution process is qualitatively different when modeled from a static network perspective than when opinions and social ties are considered as coupled dynamic systems. Yet, we still have very limited empirical evidence as to whether social ties are subject to differential decay depending on whether persons share a given set of opinions (Noel and Nyhan 2011). After all if the pre-existing sharing of opinions does not impact the dynamic stability of social ties, researchers may be correct in sticking with more tractable models that conceive of the network structure as static (and thus exogenous). In addition if persons select friends based on the fact that they share an opinion, rather than changing their opinions after they select their friends, then the conditions under which an entire network may be captured by a single opinion (under the static model) change dramatically.

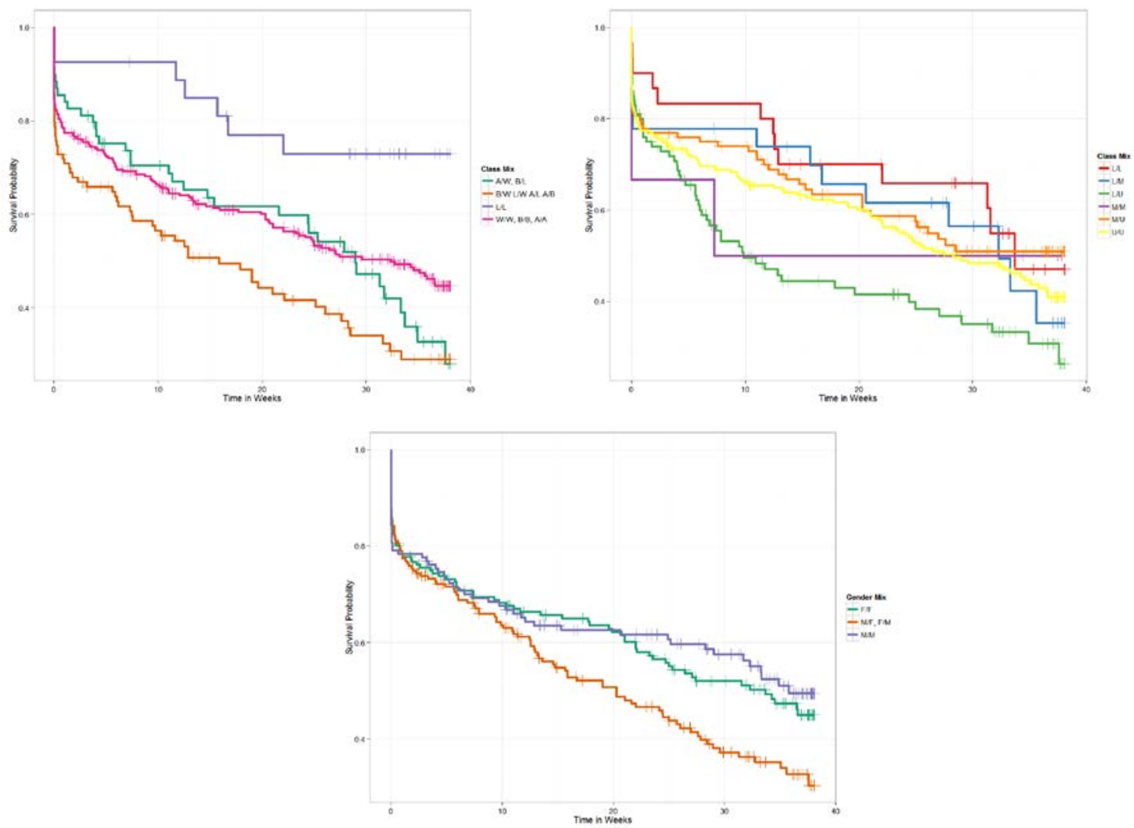

Figure 4.7. Survival Curves for Four Sets of Dyads Classified According to Ethnoracial Pairing (Left), Social Class Pairing (Right) and Gender Pairing (Bottom). Estimates of hazard rates (the probability of dissolution at a given time for each dyad type) and survivor probabilities (the probability that a tie will still be active at a given time) are computed using the Kaplan-Meier estimator in which at each time point t a risk set is specified as those objects which have survived up until t and an event set as those objects which experience an event in the interval from t to Δt. Vertical hash marks indicate right-censored cases (ties active at the end of the observation period). (Left) Ethnoracial categories are White (W), Asian (A), Black (B) and Latino (L). As shown, Latino/Latino (LL) pairs have the slowest decay rate, followed by White/White (W/W), Black/Black (B/B), Asian/Asian (AA), Asian/White (AW) and Black/Latino (B/L) pairs. Black/White (B/W), Latino/White (LW) , Asian/Black (A/B) and Asian/Latino (A/L) have the fastest decay rates. (Right) Social Class categories are "Upper Class" (Both parents college degree) (U), "Middle Class" (only one parent college degree) (M) and "Lower Class" (neither parent college degree) (L). The results show that dyads that match upper and lower class persons (L/U) decay faster than other types of dyads. (Bottom) Gender categories are Males (M) and Females (F). The results show that same gender dyads (F/F and M/M) decay at a slower rate than cross-gender dyads (M/F or F/M).

As we noted above, each CDR in the *NetSense* data contains the phone numbers of the sender and receiver along with a timestamp indicating when the event occurred. While we have thousands of CDRs for texts and calls involving a NetSense subject and people outside the study, we use only the CDRs in which both parties are in the NetSense study, because for both parties involved in the communication we have data derived from surveys on their socio-demographic characteristics, tastes, and opinions as well as other variables of interest derived from a survey administered to all NetSense subjects prior to their arrival on campus. From these CDRs we identify 505 ties formed in the first two semesters (40 weeks) among 175 students who remained in the study throughout the first year and for which we have complete data on gender, race, and parental education. We selected these three socio-demographic characteristics because sociological research shows that opinions tend to cluster within same race, gender, and class groupings (Carley 1991; Noel and Nyhan 2011). Therefore, persons of the same race, gender, or class are more likely to share the same opinions, while cross-gender, cross-racial and cross-class dyads are more likely to have different opinion. Differential decay of dyad type based on the race, class, and gender mix thus indicates that the network structure is not static but that it co-evolves with the opinion structure.

The results are shown in Fig. 4.7. We find positive evidence of network/opinion co-evolution. As shown in Fig. 4.7 (Left), cross-race dyads decay at a faster rate than same-race dyads, with Latino/Latino pairings exhibiting remarkable dynamic durability. Within the cross-race group, some pairings are more durable than others. In particular Asian/White pairings and Latino/Black pairings decay at a slower rate than other pairs. This is consistent with sociological research that reveals higher levels of agreement among members of these racial categories (Bonilla-Silva 2004). In addition, as shown by Fig. 4.7 (Right), cross-class dyads, especially those that feature a lower class person paired with an upper class person are less durable than other types of dyads, especially dyads that pair two persons of lower class origin. Sociological research shows that there is more opinion agreement within classes (especially the working class) than there is between the upper and lower classes (Weeden and Grusky 2005). This result is thus consistent with the opinion/network co-evolution thesis: ties between groups with less opinion consensus are less durable than ties within groups with high opinion consensus. This mirrors the fictional social structure of Mega-city Xanadu which is roughly divided into a three part class structure, separating the long-standing residents at the top, from the first wave of Bursuka immigrants in the model, and the more recent wave of North Razini immigrants at the bottom. Applying the results of our statistical model from the *NetSense* data to this fictional example, our results would thus predict that ties are more likely to persistent (conditional on formation) between native Cawailans and Bursuks and then they are between Razinis and natives. Finally, Fig. 4.7 (Bottom) shows that cross-gender pairings are less dynamically less stable than same gender pairings. Given the large amount of evidence for systematic opinion differences between men and women (Bolzendahl, and Myers 2004), this result is consistent with a model in which persons terminate social connections featuring opinion disagreement at a faster rate than social network connections featuring opinion agreement (Noel, and B. Nyhan 2011).

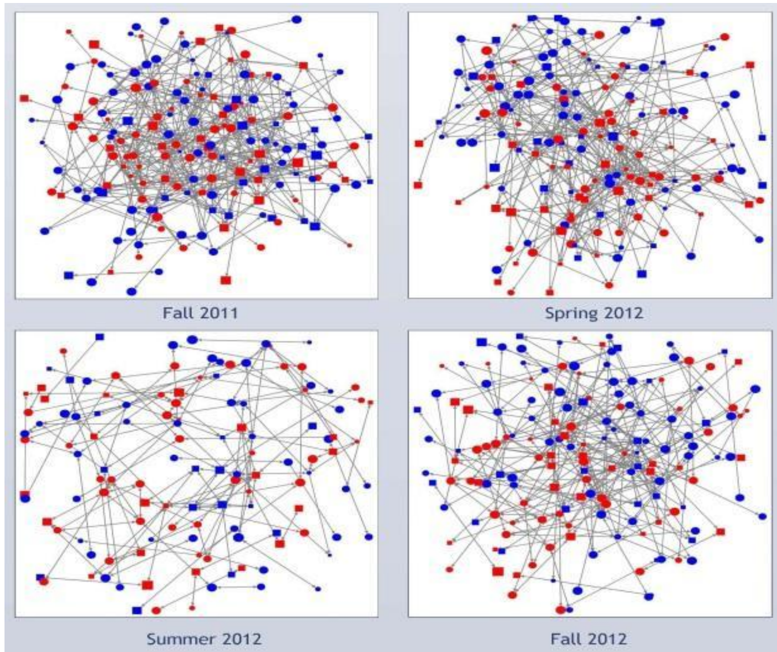

Figure 4.8. Evolution of the political opinion network among NetSense study participants (Left). Red nodes indicate conservative respondents and Blue nodes indicate Liberal respondents based on their self-placement on a seven point scale. Two participants are connected if one sends an SMS or a voice call to another who reciprocates during that time period. Participants were surveyed four times in each time period. We used stochastic actor based models (Snijders, Van de Bunt, and Steglich 2010) to model the effect of opinion similarity and network connectivity on the probability of keeping a tie. We find that for the most part, self-selection of like-minded participants and exogenous opinion accounted for the majority of overtime correlation (77.4%) across different snapshots, with opinion change via influence accounting for only a small portion of the correlation (3.9%).

Modern social systems are not only beset by divisions based on external or physical features (such as gender and race). People also sort themselves into groups and condition their social interactions based on ideas (such as religious or political ideologies). For instance, in Xanadu Bursuks have become a religious and cultural minority. This may affect their capacity to integrate and interact with others who share a different belief. Alternatively if beliefs change depending on who you interact with then the situation does not seem so hopeless. So the question to answer is: Do people select ties based on ideological similarity or does ideological similarity emerge after people form ties? In the first case, we should expect the durable existence of ideological enclaves. In the latter case, we could conceivable break these enclaves by inducing interaction across ideological divisions.

In the *NetSense* project we set out to test these hypotheses using the case of liberal/conservative placement. Fig. 4.8 summarizes results reported in (Wang, Hachen, and Lizardo 2013) based on a study of the coevolution of liberal/conservative opinions among NetSense study participants. We set out to test the hypothesis of whether the overtime correlation between personal opinions and network ties can be accounted for by self-selection of persons into dyads with like-minded others (opinion homophily) or opinion change via influence after people with different opinions became tied. We used stochastic actor based models as this simulation-based technique allows us to separate the effect of self-selection from opinion change via influence (Snijders, Van de Bunt, and Steglich 2010). We found that for the most part persons remained fairly stable in their opinions and that they selected friends based on opinion similarity rather than changing their political opinion based on the influence of people with different opinions. This provides further evidence of another mechanism (self-selection) that may help to account for the co-evolution of the opinion distribution and the social network configuration.

What lessons do these results have for our understanding of the way that social, cultural, racial, and other divisions structure social systems? First, we should not be overly sanguine or optimistic about opinion change campaigns. Just as in the Mega-city of Xanadu, we can see that persons structure their interactions based on durable memberships in social groups. These groups could be “elective” (e.g. based on ideological or membership choices) or they could be non-elective (based on traits such as gender and race). Once network dynamics are coupled to group membership, a variety of self-reinforcing mechanisms serving to segregate interaction with like-minded and socio-culturally similar others are set off. These mechanisms reify social boundaries and create strong correlations between statuses (e.g. North Razinian descent and lower social class, or Bursuka heritage and religious ideology) were initially they were only weak ones. These not only shape the formation of social networks but also their overtime evolution. Thus, even if social systems are expected to increase in terms of size and scale into the future, the underlying dynamics governing the process of tie formation, tie decay, opinion adoption, and opinion change in micro-level social systems will continue to be of primary importance.

References

Aral, S., L. Muchnik, and A. Sundararajan. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc Natl Acad Sci* USA 106(51): 21544-21549.

Baronchelli, A., M. Felici, E. Caglioti, V. Loreto, and L. Steels. 2006. Sharp transition towards shared vocabularies in multi-agent systems. *J. Stat Mech: Theory Exp* P06014.

Baronchelli, A., V. Loreto, and L. Steels. 2008. In-depth analysis of the naming game dynamics: The homogeneous mixing case *Int. J. Mod. Phys. C* 19:785-812.

Bolzendahl, C. I. and D. J. Myers. 2004. Feminist attitudes and support for gender equality: Opinion change in women and men, 1974–1998. *Social Forces* 83(2):759-789.

Bonilla-S., E. 2004. From bi-racial to tri-racial: Towards a new system of racial stratification in the USA. *Ethnic and Racial Studies* 27:6:931-950.

Carley, K. 1991. A theory of group stability. *Am. Sociol. Rev.* 56: 331-354.

Castelló, X., A. Baronchelli, and V. Loreto. 2009. Consensus and ordering in language dynamics. *Eur. Phys. J. B* 71:557–564.

Centola, D., V.M. Eguíluz, M.W. Macy. 2007. Cascade dynamics of complex propagation. *Physica A* 374: 449-456.

Centola D. 2010. The Spread of Behavior in an Online Social Network Experiment. *Science* 329(5996):1194–1197.

Dall'Asta, L. A. Baronchelli, A. Barrat, V. Loreto. 2006. Non-equilibrium dynamics of language games on complex networks. *Phys Rev E* 74:036105.

Doyle, C., Sreenivasan, S., Szymanski, B. K., and Korniss, G. 2014. Social consensus and tipping points with opinion inertia. *Physica A* 443:316–323.

Galehouse, D., Nguyen, T., and Sreenivasan, S. 2014. Impact of network connectivity and agent commitment on spread of opinions in social networks. *AHFE 2014:*2318–2329.

Gleeson J. P., and D. J. Cahalane. 2007. Seed size strongly affects cascades on random networks. *Phys Rev E*. 75:056103.

Granovetter, M. 1978. Threshold models of collective behavior. *Am. J. Sociol.* 83:1420.

Karampourniotis P. D., S. Sreenivasan, B. K. Szymanski, G. Korniss. 2015. The Impact of Heterogeneous Thresholds on Social Contagion with Multiple Initiators. *PLoS One* 10(11):143020

Morone, F., H. Makse. 2015. Influence maximization in complex networks through optimal percolation. *Nature* 524:65–68.

Latane, B. and T. L'Herrou. 1996. Spatial clustering in the conformity game: dynamic social impact electronic groups. *J. Pers. Soc. Psychol*. 70:1218.

Lu, Q., G. Korniss, and B. Szymanski. 2009. The Naming Game in Social Networks: Community Formation and Consensus Engineering. *J. Econ. Interact. Coord.* 4:221-235.

Newman, M. E. J. 2003. Mixing patterns in networks. *Phys. Rev. E* 67(2):026126.

Noel, H. and B. Nyhan. 2011. The "unfriending" Problem: The Consequences of Homophily in Friendship Retention for Causal Estimates of Social Influence. *Soc. Networks* 33(3):211-8.

Robins, G., P. Pattison., Y. Kalish and D. Lusher. 2007. An introduction to exponential random graph (p*) models for social networks. *Soc. Networks* 29(2): 173-191.

Singh, P., S. Sreenivasan, B. K. Szymanski, and G. Korniss. 2013. Threshold-limited spreading in social networks with multiple initiators. *Sci. Rep*. 3:2330.

SNAP: *Stanford Network Analysis Project*. http://snap.stanford.edu/data (Accessed: 04/23/2015).

Snijders, T. A. B., G. G. Van de Bunt, and C. E. G. Steglich, 2010. Introduction to stochastic actor-based models for network dynamics. *Soc. Networks* 32(1):44-60.

Steels, L. 1995. A self-organizing spatial vocabulary. *Artif. Life* 2:319-332.

Striegel, A., S. Liu, L. Meng, et al. 2010. Lessons learned from the netsense smartphone study. *ACM SIGCOMM Comp Comm Rev.*

Wang, C., D. S. Hachen, and O. Lizardo. 2013. The Co-Evolution of Communication Networks and Drinking Behaviors. *Proc. AAAI Fall Symposium Series*.

Watts, D. J. 2002. A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci.* USA 99:5766.

Weeden, K. A. and D. B. Grusky. 2005. The Case for a New Class Map1. *Am. J. Sociology* 111(1):141-212.

Xie, J., Sreenivasan, S., Korniss, G., Zhang, W., Lim, C., and Szymanski, B. K. 2011. Social consensus through the influence of committed minorities. *Phys. Rev. E* 84:011130.

Xie, J., Emenheiser, J., Kirby, M., Sreenivasan, S., Szymanski, B. K., and Korniss, G. 2012. Evolution of opinions on social networks in the presence of competing committed groups. *PloS One* 7(3):33215.

Zhang, W., C. Lim, S. Sreenivasan, J. Xie, B.K. Szymanski and G. Korniss. 2011. Social Influencing and Associated Random Walk Models: Asymptotic Consensus Times on the

Complete Graph. *Chaos* 21:025115.