Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

BEST FRIENDS FOREVER?
MODELING THE MECHANISMS OF FRIENDSHIP NETWORK FORMATION

Ivan Puga-Gonzalez

University of Agder
Department of Global Development and Planning
Universitetsveien 19, 4630
Kristiansand, NORWAY

Kevin McCaffree

University of North Texas
Department of Sociology
1155 Union Circle #311157
Denton, TX, USA

Fount LeRon Shults

University of Agder
Center for Modeling Social Systems at NORCE
Universitetsveien 19, 4630
4604 Kristiansand, NORWAY

ABSTRACT

The formation of friendships and alliances is a ubiquitous feature of human life, and likely a crucial component of the cooperative hunting and child-rearing practices that helped our early hominin ancestors survive. Research on contemporary human beings typically finds that strong-tie social networks are fairly small, and reveals a high degree of physical (e.g., age) and social-structural (e.g., educational attainment) homophily. Yet, existing work all too often underestimates, or even ignores, the importance of abstract, symbolic homophily (such as shared identities or worldviews) as a driver of friendship formation. Here we employ agent-based modeling to identify the optimal variable weights influencing friendship formation in order to best replicate the results of existing empirical work. We include indicators of physical and social-structural homophily, in addition to symbolic homophily. Results suggest that the optimization values that best replicate existing empirical work include strong variable weightings of kinship, shared worldview, and outgroup suspicion.

1 INTRODUCTION

The study of friendship formation is fundamental to the study of human beings. Human survival has depended on the capacity for forming cooperative groups for at least the last 2,000 to 10,000 generations. Researchers have plausibly argued that cognitive mechanisms supporting this capacity were naturally selected because they facilitated alliance formation not only with genetic kin, but also with those who share a worldview and have similar perceptions of threat (Richerson and Boyd 2005). While many animals form alliances on the basis of genetic kinship, shared short-term interests, familiarity or physical similarity, humans seem to be unique, at least by degree, in their capacity (and tendency) to form ties with conspecifics on the basis of shared worldviews and associated identities (Moffett 2019).

Existing anthropological and archaeological work on human foragers and ancestral hominins suggests that the coalitions formed from trade relationships, wartime alliances, or group inter-marriage would have been crucial for the survival and maintenance of collectives (Kelly 2013; Kelly 2019). Any model of friendship formation, then, should take account not only of the physical similarity or familiarity between individuals but also of their abstract, symbolic similarity in terms of identity or worldview (Grindal and Trettevik 2019).

In this paper, we present an agent-based model of friendship networks grounded in the existing empirical research literature on friendship formation. Our goal is to better understand what mechanisms might be influential in the formation of friendships as well as how such modeling might inform (and potentially advance) our understanding of existing empirical work. Specifically, we attempt to model findings from prior work showing that humans form friendships with those with whom they feel familiar (i.e., share roughly the same degree of educational attainment, or have a similar personality), and/or to whom they are physically similar to (i.e., same gender or same race). In addition, we model the importance of more abstract, symbolic concerns such as worldview or social attitudes.

2 METHODS

2.1 Empirical Data

We surveyed the empirical literature on networks of close friends (i.e. confidants), searching for studies reporting the average number of friends and homophily indexes (Table 1). The most common findings indicate that, on average, individuals have 3.4 close friends and that the difference between ego and friends regarding age and education years is 11.1 and 1.8 years on average. Further, friends' networks appear to be composed, on average, of 66.3%, 71.0%, 37.0%, and 89.4% of alters with the same gender, religion, family, and race as ego, respectively. We found no empirical studies reporting network composition regarding political ideology, marital status, or parental status.

Table 1: A) average number of friends; B) average difference in years relative to ego; and C) percentage of friends sharing the same trait. [1]Average of two time periods; [2]Percentage belonging to same category (i.e., low, middle, or high education); [3]Percentage belonging to age category. In **bold**, average of each category.

A)	B)		C)				Source
Friends	Age	Edu	Gender	Rel	Kin	Race	
2.8	11.2	2.0	56.7%	71.0%	NA	90.2%	(Smith et al. 2014)
3.0[1]	11.4	49%[2]	60%	NA	33.7%	93.0%	(Farkas and Lindberg 2015)
5.1	NA	NA	NA	NA	NA	NA	(Small et al. 2015)
2.8	2=40%[3] >2=35%[3] <2=25%[3]	NA	80.6%	NA	26.4%	83.4%	(Laakasuo et al. 2016)
5.0	NA	NA	NA	NA	NA	NA	(Molho et al. 2016)
2.5	10.4	1.5	68%		50.5%	91.0%	(McPherson et al. 2006)
3.4	**11.1**	**1.8**	**66.3%**	**71.0%**	**37.0%**	**89.4%**	

2.2 The Model

The BFF model described here is an extension of the SETI model (Puga-Gonzalez et al. 2019a) and the CRED model (Puga-Gonzalez et al. 2019b), both of which were written in AnyLogic, v8.5.2. Both of these models have been outlined in more detail elsewhere (Shults et al. 2018; Puga-Gonzalez et al. 2019b). Here we provide a brief description of the main features and procedures relevant for this study, i.e. those affecting the selection of friends and establishment of friends' networks.

The artificial society represented in the model is inhabited by individual human *agents* who have an education, are (un)employed, married (single), and have (or not) kids. They are categorized in relation to basic demographic variables (age, majority/minority group, education, employment, etc.); to personality (the HEXACO factors); to worldview (WV; from secular to religious [0,1]); to political ideology (PI; from left to right [0,1]); to out-group suspicion (OGS); and to affiliation (motivation to join a club, MTC, and *Frustration*, see clubs below).

On initialization, 1,000 adult agents (age 18-90) are created. These agents are assigned variables drawn from suitable distributions that may vary according to the group they belong to (majority or minority). These distributions are calibrated so socio-demographic variables that are representative of typical Western societies. The age of the agents belonging to the majority or minority group is assigned according to the age census distribution of 2011 for white British and Arab population in the UK. Years of education are drawn from a Poisson distribution with an average of 15 and 16 years for the majority and minority respectively. If at initialization the age of the agent is lower than it's years of education plus 6, its education value is then set to its age minus 6 and the occupation status is set to student. The likelihood of being employed depends on the agent's occupation status being non-student, and on its sex and group category. Males and females from the majority have a 95% and 75% chance of being employed, respectively, while minority males and females have an 85% and 65% chance. Values of the agents' HEXACO personality traits, PI, and OGS are drawn from normal distributions ($\mu = 0.5$, sd = 0.2) truncated on the range [0,1]. Other agent variables are determined by personality traits: openness determines WV, and emotionality and extraversion influence motivation to join a club (MJC). Agents get married if their age is above their marriage-age threshold. The marriage-age threshold is drawn from a Poisson distribution with average age of marriage 33.9 and 26.9 for the majority and minority group, respectively. To get married, agents must satisfy age, education, worldview, and political ideology compatibility conditions (Shults et al. 2018). Once married, couples may have a certain number of children (min = 0; max = 4; mean = 2.5). Newly born agents inherit the HEXACO personality traits (according to a correlation range), PI, and OGS from their parents (average of mother and father). WV and MJC are derived from their HEXACO personality values. Note that the process of getting married and having kids may increase the final population size to 1,500 agents.

Clubs are membership organizations that exist to support agents' WV and PI. Clubs have three different variables which determine their type: a WV variable with values *secular* or *religious*; a Cost variable with values *high* and *low*; and a PI variable with values *apolitical*, *right*, *left*. Based on the combination of these three variables, ten types of clubs were defined (Table 2).

Table 2: Types of clubs.

Eight Club types	**CODE**
Secular / low cost / PI Right	CT 1
Secular / low cost / PI Left	CT 2
Secular/ high cost / PI Right	CT 3
Secular/ high cost / PI Left	CT 4
Religious/ low cost / PI Right	CT 5
Religious/ low cost / PI Left	CT 6
Religious/ high cost / PI Right	CT 7
Religious/ high cost / PI Left	CT 8
Religious/low cost/apolitical	CT 9
Secular/low cost/apolitical	CT10

The type of club an agent may join depends on the agent's variables WV and PI. WV determines whether the agent joins a secular or a religious club. PI defines the cost and political ideology of the club. Agents with an extreme political ideology prefer high cost clubs and agents with a moderate ideology prefer low cost clubs (Figure 1). An agent joins a club when its *Frustration* value is above the threshold of its motivation to join (MTJ) a club variable. If this is the case, the agent looks for a club type matching its WV and PI values (Figure 1). Frustration and MTJ parameters are calibrated so that 70% of agents in the population are affiliated to a club and 30% are unaffiliated.

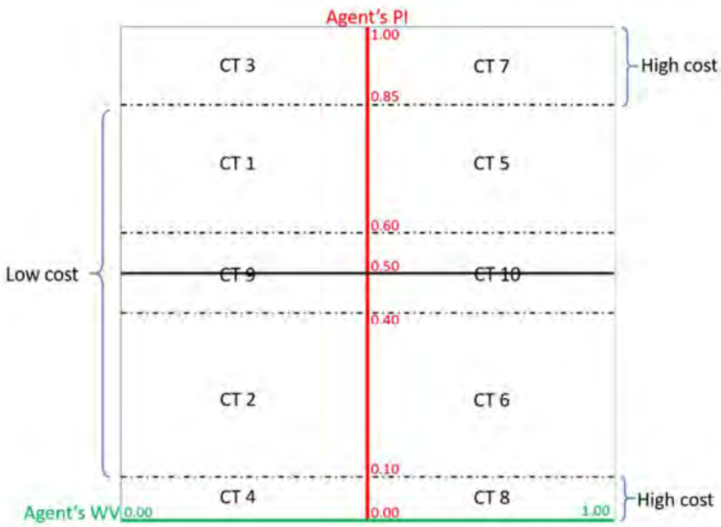

Figure 1: Club type (CT) chosen by an agent according to the agent's WV and PI values.

The maximum theoretical number of friends an agent can have in its *friends network* is determined at initialization. This number is set by drawing values from a truncated Poisson distribution with a minimum value of one and a lambda value ≥ three. Lambda is adjusted according to the agents' personality traits since it is known that individuals with high extraversion and openness have on average a higher number of friends (Molho et al. 2016). Hence, the average (lambda) number of friends an agent has is a function of these personality traits as in equation (1):

$$lambda = 3 + \left(1 * \frac{Extraversion+Openne}{2}\right) \qquad (1)$$

Thus, when an agent's extraversion and openness are both maximum (i.e., one), it will have on average one friend more than those with the lowest values of extraversion and openness. Note, however, that in practice agents may not reach their maximal number of friends. This is because for an agent to be added as a friend the friend's compatibility index (FCI) should be above a given threshold (see below). If this threshold is too high, most dyads of agents will be incompatible as friends and thus fewer than the maximum theoretical number of friends will be added as friends to the agent's friends network.

The FCI is calculated by considering nine agent variables: age, gender, relative, marital status, parental status, education, worldview, personality, political ideology and outgroup suspicion (see Table 3). When calculating the FCI, each of these variables may be weighted differently, see equation (2):

$$s^2 = \frac{\sum_i^n V_i * W_i}{\sum_i^n W_i} \qquad (2)$$

Where V_i represents the compatibility value of variable i (Table 3) and W_i its weight. The compatibility values (V_i) of each variable are determined according to the formulas in Table 3. Note that the compatibility values of OGS depend on whether or not the agents belong to the same majority/minority group (Table 3). When they belong to the same group, the value is 1 minus the absolute difference between the agents' OGS values. When agents belong to different groups, the formulas determining the compatibility values change depending on whether agents are suspicious of each other (both OGS > 0.5), not suspicious of each other (both OGS < 0.5) or one agent is suspicious (OGS > 0.5) but not the other (OGS <0.5) (Table 3).

Table 3: Formulas to determine the value (V_i) of each variable in the FCI equation.

Variable	Formula for value (V_i)	Notes
Age	Value = 1 – abs (EGO – ALTER) / AR	If value < 0, value reset to 0 AR optimized via experiments
Gender			Fem	Mal
Fem	1	0.4
Mal	0.4	1		Same gender agents more compatible than otherwise. Fem = female; Mal = male
Kin	Value = 1	Always one no matter the relative (father/mother/children).
Marital status			Married	Not married
Married	1	0.25
Not Married	0.25	1		Same marital status more compatible than otherwise
Parental status			Kid	No Kid
Kid	1	0.25
No Kid	0.25	1		Agents with(out) a kid(s) are more compatible than otherwise
Education			<HS	Co	Gr
<HS	1	0.25	0.12
Co	0.25	1	0.5
Gr	0.12	0.5	1		Compatibility decreases with higher education differences. HS: High school; Co: College Gr: Graduate
WV	Value = 1 – abs(Ego – Alter)	If abs = 0, value is 1 If abs = 1, value is 0
Personality	Value = 1 – abs(Ego – Alter)	
PI	Value = 1 – abs(Ego – Alter)	
OGS same group	Value = 1 – AbsDiff	AbsDiff = abs (OGS_{ego} – OGS_{alter})
OGS different groups		Ego/Alter	OGS <0.5	OGS >0.5
OGS <0.5	1 – (2 * AbsDiff)	1 – (3 * AbsDiff)
OGS >0.5	1 – (3 * AbsDiff)	0		**Ego and alter OGS < 0.5:** If AbsDiff = 0, value is 1 If AbsDiff >= 0.5, value is 0 **One OGS < 0.5 and other > 0.5**: If AbsDiff 0, value is 1 If AbsDiff >= 0.33, value is 0 **Ego and alter OGS > 0.5:** Value is always 0

Agents add others to their friends networks if the FCI in (2) is above a certain threshold. If the network is already full, i.e. the number of friends is equal to the maximum number of friends in (1), but the FCI of a given agent is higher than that of a current friend, then this agent has a 10% probability of replacing the friend with lowest FCI. In other words, once an agent has been added to the friends network, it is not so easily replaced. Friendships are unidirectional; i.e., when ego adds an agent to its friendships network, ego considers this agent as a friend, but the reverse may not be true. Hence, once ego has added an agent to its friends network, the newly added agent is given the chance to add ego as a friend. However, whether ego is added as a friend depends on whether alter's friendships network is not full. If so, then if the FCI is higher than that of the friend with the lowest FCI, this friend is replaced with a 10% probability.

2.3 Simulations, Parameters Variation and Data Collection

We first ran optimization experiments, using the AnyLogic optimization engine, to find the optimal combination of weight values so that the composition of the friends networks emerging in the model fit approximately with previously observed empirical values (Table 1). We constrained the potential weight values in the range [1,20]. Thus, at maximum, a variable could be weighted 20 times higher than another. In addition to the values of the weights (W_i) of the nine variables previously mentioned, we also optimized the value of the age range (AR) variable (Table 3). This variable determined the age range within which individuals could be considered as candidates for being friends. The potential values of this variable AR were also in the range [1,20]. Optimization experiments were run for 52 time-steps. We considered each time step to represent a week, so the experiments were run for a *year*. Every time step agents sample up to 20 agents randomly drawn from each (when available) of their social networks (club, work, neighborhood, offline, online, family, and friends of friends networks). Thus, agents could add (or replace) friends to (or from) their networks for a year. The friendship threshold was set to 0.6. We ran 10 optimization experiments from which we obtained 10 different combinations of weighted values and age ranges of friends.

Next, we took the average values of the weights and age range variables obtained from the optimization experiments and ran simulations fixing these values and varying the values of the FCI threshold: 0.5, 0.55, 0.6. For each of these three threshold values we ran 100 simulations and analyzed the effect on the average size and composition of the friendship networks. As in the optimization experiments, simulations were run for 52 time-steps, or one year. At the end of each simulation we collected the data of every agent older than 12 years old as well as that of the agents in its friends network.

2.4 Friends Network Analyses

Network composition was calculated as follows. Age composition is the average number of age difference between ego and its friends. Education is the average number of difference in education-years between ego and its friends. Gender and group (majority/minority) composition are the percentages of friends that are the same gender and group as ego. Kin is the percentage of friends that are related to ego (father, mother, or children). Marital status and parental status are the percentages of friends with the same marital status or parental status as ego. WV and PI are the percentage of friends that share the same worldview and PI as ego. WV is defined as secular when WV < 0.5, and as religious when WV > 0.5. Similarly, PI is defined as left when PI < 0.5 and right when PI > 0.5.

Additionally, we ran a descriptive network analysis at the population level with the aggregated friends networks. We calculated the following global metrics: 1) density, 2) modularity) clustering coefficient, 3) number of clusters, 4) diameter, 5) average shortest path, and 6) percentage of reciprocal friendships.

Definition 1 Density: The proportion of the potential links in a network that are actual connections. A potential link is a connection that could exist between two agents and an actual connection is one that actually exists.

Definition 2 Modularity: The difference between the proportion of the total association of individuals within clusters and the expected proportion, given the summed associations of the different individuals. A high value means a high number of links within a cluster, but few contacts between clusters and low modularity means a homogeneous distribution of links between all agents.

Definition 3 Clustering coefficient: The proportion between the number of triads (links among three agents) in the network divided by total possible number of triads.

Definition 4 Number of clusters: The number of disconnected clusters in the overall network.

Definition 5 Diameter: The shortest distance between the two most distant agents in the network.

Definition 6 Average shortest path: The number of average links needed to communicate between any two given agents in the network.

Definition 7 Percentage of reciprocal friendships: Percentage of connected agents (dyads) that consider each other as a friend.

3 RESULTS

3.1 Optimization Experiments

We ran a total of 10 optimization experiments. From these, we obtained 10 different combinations of weighted values for each of the variables in the FCI equation and age range of friends (Table 3). The mean [min-max] values are shown in Table 4. The optimized weighted values show that to obtain networks with a composition like those observed in empirical data (Table 1), agents must give the highest weights to relatives, and those with a similar worldview and outgroup suspicion. Thereafter, age, personality, and gender seem to be most important. Finally, the traits with the lowest importance were marital status, parental status, education, and political ideology. The age range value shows that agents may be considered as potential friends for other agents with an age difference of up to 14 years.

Table 4: Average, minimum, and maximum weighted values for each variable in the FCI obtained by running 10 optimization experiments.

	Average	**Min**	**Max**
Kin	16.2	15.5	18.0
Worldview	9.5	9.5	9.7
Out Group	8.0	6.5	10.2
Age	5	5.0	5.1
Personality	3.8	3.5	4.4
Gender	3.6	3.5	3.9
Marital status	1.1	1.0	1.7
Parental status	1	1.0	1.1
Education	1	1.0	1.3
Political Ideology	1	1.0	1.2
Age Range	14.1	10.0	20.0

3.2 Simulations and Network Composition

Table 5 shows the average values (of 100 simulations) for network size and friends network composition, for different FCI values, compared to the observed empirical values (Table 1). All averages of network sizes for the different FCI fall within the observed empirical values (Tables 1 and 5). However, network size appears to increase with decreasing values of FCI. This is obvious since the less selective agents are when choosing friends, the easier it is for them to accept others into their friendship networks. This also means that at FCI values of 0.6, the number of friends an agent has is slightly lower than the theoretical maximum given by (1), and that higher values of FCI will result in network sizes much lower than observed in empirical data. Regarding network composition traits, gender, group, and kin fall within the range observed in empirical networks, for the three values of FCI (4-6 in Table 5). Age composition is slightly lower than the minimum observed for CFI values of 0.50 and 0.55 (2 in Table 5). Education is slightly higher than the maximum observed value for all three values of FCI (3 in Table 5); and Worldview is also slightly higher than the maximum observed for FCI values of 0.50 and 0.55 (7 in Table 5). Overall, however, we can see that the combination of the weight values (Table 3) can accurately reproduce the network composition observed in empirical data at the three different values of FCI. We also observed that the simulated friends networks are composed of slightly more friends sharing the same marital and parental status as ego (8-9 in Table 5). Regarding political ideology, friends networks are composed of an average of 63% of friends sharing the same ideology as ego (10 in Table 5). Although we lack empirical data to compare network composition related to these last three traits, these findings from the BFF model can generate new hypotheses about the establishment and maintenance of friendship networks that can be further tested using other research methods.

Table 5: Average value of 100 simulations [1^{st}-3^{rd} quartile] for friendship network size and composition; and min-max values observed from empirical networks (see Table 1).

	FCI 0.50	**FCI 0.55**	**FCI 0.60**	**Observed value [Min-Max]**
	Average [1^{st}-3^{rd}]	**Average [1^{st}-3^{rd}]**	**Average [1^{st}-3^{rd}]**	
1) Net Size	3.5 [3.5-3.5]	3.4 [3.4-3.5]	2.3 [2.3-2.3]	2.08-5.27
2) Age	9.9 [9.6-10.1]	9.6 [9.4-9.9]	11.6 [11.1-12.0]	10.35-11.4
3) Education	2.4 [2.3-2.4]	2.3 [2.3-2.4]	2.2 [2.2-2.3]	1.48-2.04
4) Gender	70.2 [69.6-70.9]	70.6 [70.0-71.3]	64.7 [64.0-65.6]	56.7-80.6
5) Group	91.4 [91.0-91.8]	91.7 [91.3-92.2]	92.8 [92.3-93.3]	83.4-93.0
6) Kin	25.5 [24.7-26.3]	27.9 [27.2-28.5]	39.4 [38.4-40.3]	26.4-51.0
7) WV	74.5 [73.6-75.3]	74.5 [73.7-75.2]	71.0 [70.0-72.0]	71.0
8) Parental status	53.0 [52.1-53.7]	52.1 [51.3-53.1]	48.3 [47.4-49.0]	NA
9) Marital status	53.9 [53.1-54.7]	53.0 [52.1-53.7]	49.5 [48.7-50.3]	NA
10) PI	63.0 [62.3-63.6]	63.2 [62.7-64.0]	64.0 [63.0-65.0]	NA

3.3 Aggregated Networks' Characteristics

The networks from simulations with FCI values of 0.60 were smaller and less densely connected than those with an FCI of 0.50 and 0.55 (2-3 Table 6). A probable explanation for this result is that when the FCI is equal to 0.60, the percentage of agents that have no friends is 12%, whereas this percentage is almost zero for 0.50 and 0.55 FCI values (1 in Table 6). Since agents that have no friends cannot be included in the aggregated networks, the size and density of the networks are smaller. Surprisingly, the clustering coefficient was higher in networks with FCI values of 0.60 (4 in Table 6), meaning that these networks formed more tightly linked groups of friends. This is further corroborated by the higher modularity as well as the higher percentage of reciprocal friendships at FCI values of 0.60 (5-6 in Table 6). In fact, the values of these metrics appear to decrease with decreasing values of FCI. Diameter and mean shortest path also appear to decrease with decreasing values of CFI (7-8 in Table 6). Diameter is a measurement of the number of links necessary to reach the two most peripheral agents in the network and the mean shortest path is the average number of links necessary for one agent to reach any other given agent in the network. Hence in networks with 0.60 FCI values, individuals within the clusters are more tightly connected but their connections to individuals from different clusters appear looser; thus, the values of diameter and average shortest path are higher. Finally, networks with 0.60 FCI values have a high number of unconnected clusters. In fact, these networks are composed of one big cluster and many small ones that are disconnected from the biggest clusters. In such cases, the smaller clusters vary in size from only two connected agents to up to nine connected agents. Note that if we rerun the analysis considering only the biggest cluster from networks with 0.60 FCI values, results remain qualitatively the same as those in Table 6.

4 DISCUSSION

The results of our simulation experiments suggest that the optimal variable weights for producing results approximating real-world friendship networks are highest for kinship, followed by worldview and out-group suspicion. Optimal weights for age, personality, and gender were relatively lower, while weights for variables such as marital status, education, and political ideology were lower still.

Table 6: Average global network metric values of 100 friends networks at the population level, [1st – 3rd] quartile, at different values of FCI.

	FCI 0.50	FCI 0.55	FCI 0.60
	Average 1st-3rd	**Average 1st-3rd**	**Average 1st-3rd**
1) % of agents with no friends	0.0 [0.0-0.0]	0.5 [0.3-0.6]	12.2 [11.4-13.0]
2) Network size	1296 [1282-1311]	1286 [1273-1299]	1130 [1109-1148]
3) Density	0.44 [0.43-0.44]	0.41 [0.40-0.41]	0.29 [0.28-0.29]
4) Clustering coefficient	0.12 [0.11-0.12]	0.16 [0.16-0.16]	0.25 [0.25-0.26]
5) Modularity	0.59 [0.59-0.60]	0.65 [0.64-0.65]	0.81 [0.81-0.82]
6) Reciprocity	37.7 [37.0-38.3]	48.4 [47.6-49.2]	76.4 [75.7-77.0]
7) Diameter	9.7 [9.0-10.0]	11.4 [11.0-12.0]	20.6 [19.0-22.0]
8) Mean shortest path	4.8 [4.8-4.8]	5.3 [5.3-5.3]	8.2 [8.0-8.4]
9) Number of clusters	1.0 [1.0-1.0]	1.6 [1.0-2.0]	28.2 [25.0-32.0]

The outsized importance of kinship in friendship formation should not surprise us. Individuals are exposed to kin by virtue of their family membership, making kin very accessible as interaction partners. This is consistent with existing empirical work showing that those who come from large, extended families form fewer non-kin friends (Dunbar 2014). A high weighting placed on kinship is also consistent with an inclusive-fitness kin selection model (Hamilton 1964) of friendship. The high weighting that our simulated agents placed on kinship-ties, especially in the formation of strong-tie networks (the target of our BFF model) is therefore consistent with existing theory.

However, it is also possible that the high weighting of kin in our model is something of an artifact of the model. It may have resulted from the fact that some agents in our model have no kin, thus requiring other agents to weight kinship more highly in order to compensate for those who have none. If other agents with kin did not weight kinship highly in their formation of friendships, the model results might not have approximated the results seen in real world studies. This could be an artifact of the model since it is unlikely that real people form more friendships with kin to compensate for others in their neighborhood who have no kinship ties. Nevertheless, the high weighting of kinship in friendship formation found in the model is consistent with existing theory and research. Another limitation of the current implementation of the model is that it does not account for other cross-cultural ways of weighting kin, relative to non-kin (e.g., people in poorer, more agrarian, and collectivist cultures may have more kinship ties in their strong-tie networks than those in richer, more individualistic, and urban cultures).

The high weighting of shared worldviews and scores on out-group suspicion is also consistent with existing theory, in particular, with theories of friendship formation that highlight the evolutionary importance of coalitions formed with non-kin (Richerson and Boyd 2001; Henrich 2017). Assessing whether or not a close ally shares one's general view about the social landscape, and whether or not an ally agrees about the threats posed (or not posed) by other groups, would have been central to the formation and maintenance of group boundaries and in-group cooperation (Moffett 2019). Certainly, not all worldviews contain strong elements of out-group suspicion, which is why we modeled these variables separately.

We operationalized worldview here in relation to a distinction between "secular" and "religious" views of the nature of reality, agreement on which is expected to lead to assortative friendship formation in the model. Theorists speculate that early life physical dependency of human infants on others, paired with the substantial neural plasticity this affords, may have facilitated selection for cognitive mechanisms facilitating the rapid ascertainment of emotion-laden group norms (Henrich 2017; Heyes 2018). These group norms would have been important for maintaining the in-group cooperation necessary for collective hunting, resource acquisition, and child-rearing. Moreover, cross-cultural research suggests that these group norms tend to be rooted in either supernaturalist religious or naturalist secular worldviews (Flannery and Marcus 2012).

The results of our simulation experiments also reveal that higher weightings of age, personality, and gender play a more important role in friendship network formation in comparison to (for example) education, political ideology, parental status, and marital status. The low weighting of these latter variables in our optimization experiments is harder to explain. It is possible that the association between worldview and club membership is too strong in our models, and is consequently biasing the weighting of variables. In the real world, occupations, neighborhoods, and clubs are often stratified by educational attainment, income, or marital status – all three of which tend to be correlated (Blau and Schwartz 1984). In the present model, work, neighborhood, and club networks were not structured by these variables. This is an important limitation of our model because it means that we were not able to capture the "induced network homophily" that results from individuals tending to work in jobs, live in neighborhoods, or join clubs in specific geographic areas, which could reveal higher (or lower) levels of educational attainment, income, marriage, or fertility. The clustering of people with certain levels of educational attainment or of a certain marital status into particular neighborhoods or clubs should constrain the possible friendships that are likely to form (McPherson and Smith-Lovin 1987). Put differently, our model can account for friendships formed as a result of choice, but not as a result of constraints in exposure. Not all friendships form as a result of specific choices made by individuals; friendships also form because of constraints on opportunities to meet different types of people. This could account for why the weightings on age or gender, and especially education and marital status, are fairly low in our simulation results.

The low weighting of political ideology might seem even more surprising given the discussion above on the importance of worldview and identity for "fictive kin" or close friendship formation. It is likely that the low relevance of political ideology may be an artifact of the algorithm we used for determining which agents will join particular clubs. In the present model, agents are programmed to join clubs that match their worldview as well as their political ideology. The weighting placed on worldview, then, may in an important sense "wash out" any need for agents to weight political ideology – friends selected on the basis of shared worldviews are already likely to also share ego's political ideology. To account for the auto-correlation of club, worldview, and political ideology, future models might more sharply distinguish when individuals join clubs based on their worldview, but not based on political ideology.

On the other hand, future work might treat worldview as a proxy for political ideology instead of being modeled distinctly; people may not be making fine-grained distinctions about political ideology among their friends once they are confident that they share their worldviews. Indeed, existing empirical work suggests that political ideologies might be derivative of views of reality, human nature, and society (Lakoff 1996) which we most closely approximate here with our "worldviews" variable. For example, religious conservatives are more likely to construe human nature (and their politics) in terms of conflict and competition between groups, whereas secular liberals are more likely to construe human nature (and their politics) in terms of cooperation between relatively malleable individuals (Pinker 2002). It may be, then, that what most distinguishes political ideologies has to do with worldview, and that the latter is sufficient for driving friendship formation. Further work is required to answer such research questions.

Finally, our findings may also have important implications for current research on the application of social network simulation to societal challenges such as preventing the spread of misinformation, stigma, and anxiety in the wake of a pandemic. Social simulation is increasingly being used to study the mechanisms and dynamics at work behind similar challenges related to emotional contagion (Fan et al. 2018; Zeng and Zhu 2019). Most simulated networks, however, have inadequately realistic cognitive architectures for their agents and are overly simplistic in their representation of the factors that impact their homophily. Our results might be utilized to develop artificial societies with more psychologically and sociologically realistic architectures, which could then be used for experimenting with competing policy proposals for addressing such societal challenges.

ACKNOWLEDGMENTS

We thank Prof. Dr. Wesley Wildman and Prof. Dr. Ryan Cragun for constructive comments and discussion on the design of Friends Compatibility Index equation and club types.

REFERENCES

Blau, P. M., and J. E. Schwartz. 1984. *Crosscutting Social Circles: Testing a Macrostructural Theory of Intergroup Relations*. New York: Academic Press.

Dunbar, R. I. M. 2014. "The Social Brain: Psychological Underpinnings and Implications for the Structure of Organizations". *Current Directions in Psychological Science* 23(2):109-114.

Fan, R., K. Xu, and J. C. Zhao. 2018. "An Agent-Based Model for Emotion Contagion and Competition in Online Social Media". *Physica a-Statistical Mechanics and Its Applications* 495:245-259.

Farkas, G. M., and E. Lindberg. 2015. "Voluntary Associations' Impact on the Composition of Active Members' Social Networks: Not an Either/or Matter". *Sociological Forum* 30(4):1082-1105.

Flannery, K., and J. Marcus. 2012. *The Creation of Inequality: How Our Prehistoric Ancestors Set the Stage for Monarchy, Slavery, and Empire.* Cambridge, Massachusetts: Harvard University Press.

Grindal, M., and R. Trettevik. 2019. "Perceived Similarity in Identity Meanings, Identity Verification, and Positive Emotions". In *Identities in Everyday Life* edited by J. E. Stets and R. Serpe, 35-52. New York: Oxford Press.

Hamilton, W. D. 1964. "The Genetical Evolution of Social Behaviour. I". *Journal of Theoretical Biology* 7:1-16.

Henrich, J. 2017. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton, New Jersey: Princeton University Press.

Heyes, C. 2018. *Cognitive Gadgets: The Cultural Evolution of Thinking*. Cambridge, Massachusetts: Harvard University Press.

Kelly, C. 2019. *The Fifth Beginning: What Six Million Years of Human History Can Tell Us About Our Future*. Berkeley, California: University of California Press.

Kelly, R. 2013. *The Lifeways of Hunter-Gatherers: The Foraging Spectrum.* Cambridge, United Kingdom: Cambridge University Press.

Laakasuo, M., A. Rotkirch, V. Berg, and M. Jokela. 2016. "The Company You Keep: Personality and Friendship Characteristics". *Social Psychological and Personality Science* 8(1):66-73.

Lakoff, G. 1996. *Moral Politics: How Liberals and Conservatives Think*. Chicago, Illinois: University of Chicago Press.

McPherson, J. M., and L. Smith-Lovin. 1987. "Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups". *American Sociological Review* 52(3):370-379.

McPherson, M., L. Smith-Lovin, and M. E. Brashears. 2006. "Social Isolation in America: Changes in Core Discussion Networks over Two Decades". *American Sociological Review* 71(3):353-375.

Moffett, M. W. 2019. *The Human Swarm: How Our Societies Arise, Thrive, and Fall*. New York: Basic Books.

Molho, C., S. G. B. Roberts, R. E. de Vries, and T. V. Pollet. 2016. "The Six Dimensions of Personality (Hexaco) and Their Associations with Network Layer Size and Emotional Closeness to Network Members". *Personality and Individual Differences* 99:144-148.

Pinker, S. 2002. *The Blank Slate: The Modern Denial of Human Nature*. New York: Penguin Putnam.

Puga-Gonzalez, I., D. Voas, W. J. Wildman, S. Diallo, and F. L. Shults. 2019a. "Minority Integration in a Western City: An Agent-Based Modelling Approach". In *Human Simulation: Perspectives, Insights and Applications*, edited by S. Diallo, W. J. Wildman, F. L. Shults, and A. Tolk. New York: Springer.

Puga-Gonzalez, I., J. W. Wesley, K. McCaffree, T. C. Ryan, and F. L. Shults. 2019b. "Incredulity in Artificial Societies". In *Social Simulation Conference,* September 23rd-27th, Mainz, Germany.

Richerson, P. J., and R. Boyd. 2001. "The Evolution of Subjective Commitment to Groups: A Tribal Instincts Hypothesis". In *Evolution and the Capacity for Commitment*, edited by R. M. Nesse, 186-202. New York: Sage.

Richerson, P. J., and R. Boyd. 2005. *Not by Genes Alone*. Chicago: University of Chicago press.

Shults, F. L., W. J. Wildman, S. Diallo, I. Puga-Gonzalez, and D. Voas. 2018. "The Artificial Society Analytics Platform". In *Social Simulation Conference*, August 20th-24th, Stockholm, Sweden.

Small, M. L., V. Deeds Pamphile, and P. McMahan. 2015. "How Stable Is the Core Discussion Network?". *Social Networks* 40:90-102.

Smith, J. A., M. McPherson, and L. Smith-Lovin. 2014. "Social Distance in the United States: Sex, Race, Religion, Age, and Education Homophily among Confidants, 1985 to 2004". *American Sociological Review* 79(3):432-456.

Zeng, R., and D. Zhu. 2019. "A Model and Simulation of the Emotional Contagion of Netizens in the Process of Rumor Refutation". *Scientific Reports* 9(1):14164.

AUTHOR BIOGRAPHIES

IVAN PUGA-GONZALEZ is a post-doctoral researcher at the University of Agder. His research interest is modeling and simulation of complex adaptive systems. He holds a PhD in Theoretical Biology from the University of Groningen. His email address is ivan.puga-gonzalez@uia.no.

KEVIN MCCAFFREE is an assistant professor of sociology at the University of North Texas. His research interests are sociological theory, cultural evolution, and criminology. He holds a PhD in Sociology from the University of California, Riverside. His email address is kevin.mccaffree@unt.edu.

FOUNT LERON SHULTS is professor at the Institute for Global Development and Social Planning at the University of Agder and research professor at the Center for Modeling Social Systems at NORCE in Kristiansand, Norway. He has published 18 books and over 130 scientific articles and book chapters on topics such as the ethics of computer modeling, philosophy of science, and the cognitive science of religion. His email address is leron.shults@uia.no.