
A Psychologically-Motivated Model of Opinion Change with Applications to American Politics

Peter Duggins¹

¹Computational Neuroscience Research Group, University of Waterloo, 328 Breezewood Crescent, N2L5K5 Waterloo, Ontario, Canada

Correspondence should be addressed to psipeter@gmail.com

Journal of Artificial Societies and Social Simulation 20(1) 13, 2017

Doi: 10.18564/jasss.3316 Url: <http://jasss.soc.surrey.ac.uk/20/1/13.html>

Received: 09-09-2016

Accepted: 25-11-2016

Published: 31-01-2017

Abstract: Agent-based models are versatile tools for studying how societal opinion change, including political polarization and cultural diffusion, emerges from individual behavior. This study expands agents' psychological realism using empirically-motivated rules governing interpersonal influence, commitment to previous beliefs, and conformity in social contexts. Computational experiments establish that these extensions produce three novel results: (a) sustained "strong" diversity of opinions within the population, (b) opinion subcultures, and (c) pluralistic ignorance. These phenomena arise from a combination of agents' intolerance, susceptibility and conformity, with extremist agents and social networks playing important roles. The distribution and dynamics of simulated opinions reproduce two empirical datasets on Americans' political opinions.

Keywords: Agent-Based Model, Opinion Dynamics, Social Networks, Conformity, Polarization, Extremism

Introduction

- 1.1** Opinions are mutable: individuals revise their beliefs through social interaction, personal experience, and reflection, while societal norms shift in response to global events and public opinion. Opinion change at the individual and societal scales interact to produce political polarization, cultural globalization, and other important social trends. To understand these phenomena and design appropriate interventions, we need quantitative tools that simulate the psychological and social aspects of opinion change. For example, models of interpersonal communication will help activists organize grassroots support, help leaders design effective campaigns, and help peacekeepers prevent the spread of extremism. Computational models of opinion change have studied the relationship between polarization, social influence, and political intolerance (Hegselmann et al. 2002; Deffuant et al. 2002; Jager & Amblard 2005; Baldassarri & Bearman 2007), while models of cultural diffusion have improved our understanding of cultural convergence (Carley 1991; Mark 1998), subculture formation (Bednar et al. 2010; Kitts 2006), and cultural stability within organizations (Carroll & Harrison 2002; Kitts & Trowbridge 2007).
- 1.2** Building multi-level, quantitative, predictive models of opinion change is challenging because opinions arise from a multitude of neurological, psychological, and social processes. Empirically, the extent to which people are persuaded by each others' subjective evaluations depends on numerous factors, including: previous beliefs and a desire to minimize cognitive dissonance (Petty et al. 1997); motivations to be accurate, self-consistent, and socially accepted (Wood 2000; Cialdini & Goldstein 2004); issue framing, emotional arousal, and cognitive elaboration (Gawronski & Bodenhausen 2006); self-esteem (Pool et al. 1998); social norms (Lapinski & Rimal 2005); and more. Mathematical and computational models help formally investigate both the interplay of internal psychological forces and the feedback between opinion change and social influence among many individuals. Unfortunately, models have historically neglected important elements of social psychology, assuming that individuals behave identically, rationally, or with perfect information. This raises questions about whether their results properly inform our understanding of human societies.
- 1.3** Agent-based models (ABMs) seek to explain macroscopic outcomes by showing that artificial societies populated by psychologically-plausible software individuals can, when initialized in a virtual environment and

evolved through time, endogenously “grow” complex social phenomenon (Epstein 2006). Three features of ABMs make them ideal for modeling opinion change. First, agents are autonomous and heterogeneous: each individual has distinct internal attributes, such as an intolerance of opposing views, a propensity to socially conform, or a tendency towards stubbornness. Second, agents can be psychologically and cognitively authentic, endowed with rational, emotional, and social thinking of arbitrary complexity (Epstein 2014). Third, agents interact locally in an explicitly defined space: individuals have incomplete information about the world, and interact in social networks of plausible size and composition, causing influence to spread through the population in a manner constrained by personal connections.

- 1.4** Although a rich literature of agent based opinion dynamics models already exists, several important questions remain unanswered:

1. **Why do societies maintain a diversity of opinions?** Previous models have shown that when agents exchange interpersonal influence, their opinions either converge to a single value (consensus) or diverge to homogeneous opinion groups (polarization) (Hegselmann et al. 2002; Deffuant et al. 2002; Salzarulo 2006; Jager & Amblard 2005). However, there are reasons to believe that these models cannot maintain a pluriformity of opinions: for an excellent overview see Mäs et al. (2014). Essentially, when agents positively influence one another in fully-connected networks (where each actor connects to each other by at least one influence path), all agents will converge to a homogeneous opinion (Abelson 1964). **Weak diversity**, defined as the convergence of opinions to a finite number of attractor states, can be maintained when opinion subcultures become isolated, breaking full connectivity. This outcome is common in **bounded confidence** models, in which agents who are sufficiently dissimilar cease to influence one another entirely. **Strong diversity**, defined as a smooth distribution of opinions along a continuous ideological spectrum, cannot survive in these models (Flache & Macy 2011; Mäs et al. 2010), even when the mechanisms of social influence incorporate noise and other minor deviations (Klemm et al. 2003; De Sanctis & Galla 2009). As noted by Mäs et al. (2014), “This [homogeneity] contradicts the high degree of persistent diversity that we observe in many social settings [...leading] researchers to develop extensions of the classical models to explain emergence and persistence of diversity.”
2. **Will subcultures of opinions survive in a well-connected population?** Pockets of extreme opinions exist within moderate real-world societies. Although such subcultures have emerged in previous models (Axelrod 1997; Amblard & Deffuant 2004; Carley 1991; Mark 1998), they survive only because of the previously-mentioned rules that curtail interpersonal influence. When these rules are relaxed (and others are introduced to promote strong diversity), will these groups dissolve, as occurs in cultural diffusion models with the introduction of noise and long-range interactions (Klemm et al. 2003; De Sanctis & Galla 2009)? Do other psychosocial forces explain why opinions in real societies organize according to social networks (e.g. smoking habits (Christakis & Fowler 2008))?
3. **How do hidden opinions affect societal opinion change?** The views we express in public often differ from those we hold privately, a situation that undoubtedly affects individual and societal opinion dynamics in some manner. For example, citizens living under authoritarian rule rarely express their true dissatisfaction with the government for fear of oppression (Kuran 1989). The disconnect between true and expressed opinions creates tensions that have been implicated in major social and political events such as the fall of Communism (Kuran 1997) and the Arab Spring (Goodwin 2011). Though some models of “preference falsification” exist, they assume global information and rational action, and have not yet been applied to the study of interpersonal influence within social networks.

- 1.5** In this study, I aim to answer these questions by studying the relationship between the social psychology of personal opinion change and the distributions, dynamics, and clustering of opinions across a population of agents. In Section 2, I review the literature on the social and psychological forces that drive opinion change. In Section 3, I suggest why extending previous models to include these forces will increase a model’s explanatory power and help answer the above question, then present the ISC model. In Section 4.1, I justify these increases in model complexity by showing that each additional psychological force explains a new set of behaviors outside the scope of simpler models. In Section 4.2, I present the full model results and describe the emergence of strong diversity, persistent subcultures of opinions, and pluralistic ignorance. In Section 5, I compare these results with empirical data on Americans’ political opinions. I conclude by summarizing the major findings, answering the questions posed in the introduction, and proposing a research agenda for agent models in the social sciences.

Social Psychology of Opinion Change

- 2.1** **Social influence** is a process in which the social exchange of information causes individuals to reevaluate their own opinions on a subjective issue. Arguably the most important feature of social influence is **homophily**, the principle that contact between similar people occurs more frequently and has greater impact than contact between dissimilar people. Empirical evidence for homophily and its effects on social influence abounds: for an overview, see McPherson et al. (2001). Interpersonal influence among friends is known to engender common attitudes (Friedkin 1984; Marsden 1988; Friedkin 1993), while the strength of dyadic connections concurrently increases with similarity (Carroll & Harrison 2002; Kitts et al. 1999). On the other hand, interactions can impart negative social influence if opinions differ greatly (Rosenbaum 1986; Smeaton et al. 1989), causing individuals to adopt more extreme attitudes when exposed to counterattitudinal arguments (Lord et al. 1979; Miller et al. 1993; Taber & Lodge 2006). Homophily is a cornerstone of opinion dynamics models: individuals exert social influence on each other proportional to their ideological similarity. In dyadic conversations, similarity encourages consensus, while dissimilarity fosters polarization. A lineage of bounded confidence models have shown that a population with high **tolerance** (a parameter governing the relationship between opinion similarity and the magnitude of influence) leads to consensus, while low tolerance leads to polarization (Hegselmann et al. 2002; Deffuant et al. 2002; Dandekar et al. 2013; Salzarulo 2006). More recent studies have extended this classic formalism to include distancing (negative influence weights) (Jager & Amblard 2005) and a drive for distinctiveness (random opinion perturbations) (Mäs et al. 2014); these additional bottom-up mechanisms encourage agents to adopt opinions that differ from those of their neighbors, which helps oppose the homogenizing positive influence in bounded confidence models. As a result, these studies produce new patterns of, and insights into, opinion consensus, clustering, and polarization, though these additions appear insufficient to produce strong diversity in these models.
- 2.2** Social influence does not take place in a vacuum, but in an environment filled by people who seek social acceptance and who judge each other upon personality and beliefs. **Conformity** describes an individual's desire to gain social approval and avoid rejection by expressing normative beliefs. There is substantial empirical evidence of people misrepresenting their true beliefs (Wood 2000; Cialdini & Goldstein 2004; Asch 1951), though some "anticonformists" will express non-normative beliefs so as to appear more distinct (Imhoff & Erb 2009). Together, conformity and distinctiveness lead to **pluralistic ignorance** (Prentice & Miller 1993), a condition in which the true distribution of opinions in society differs from what is spoken and heard in public. Pluralistic ignorance makes people unaware of others' true beliefs; a lack of accurate information can, through the mechanisms of social influence, feedback to change people's true opinions. Examples of this phenomenon include illusory support for the communist regime in the Soviet Union (Kuran 1997), inaccurate estimates among whites about white support for racial segregation (O'Gorman 1975), and differences between private drinking behavior and public support for drinking on college campuses (Prentice & Miller 1993). Despite current enthusiasm for studying the effects of conformity on opinion change (Jarman et al. 2015; Smaldino & Epstein 2015), ABMs have yet to investigate the repercussions of agents' verbal belief falsification on public opinion.
- 2.3** The way an individual receives and internalizes others' beliefs can be as important as the content and context of the influence. People who hold strong opinions are **committed** to their beliefs: they resist opinion change, because it would challenge their political worldview and induce cognitive dissonance, and because they judge contrary information as invalid due to confirmation bias (Ajzen 2001). Strongly opinionated individuals have been shown to reject opinions contrary to their own belief and even become more extreme. On the other hand, moderately opinionated individuals are **susceptible** to opinion change and will more readily internalize beliefs presented by others (Lord et al. 1979; Miller et al. 1993; Taber & Lodge 2006). Surprisingly, few models of opinion change have looked into how susceptibility and commitment help sustain diversity and prevent homogenization of small cultural groups (Baldassarri & Bearman 2007).
- 2.4** Finally, the social networks through which individuals interact determine how opinion change spreads through society. These networks can be characterized by statistical descriptions such as the degree of connectivity (average size of a social network); real-world networks have positive assortativity (people with large networks tend to know others with large networks), low whole-network density (most people don't know each other), and high but heterogeneous clustering. Though simulations have confirmed that the size and composition of social networks strongly affect opinion change, their outcomes vary widely with the models' assumptions about the network (Amblard & Deffuant 2004; Centola et al. 2005), which rarely take these empirical regularities into account. One procedure which does effectively reproduce these statistics is the social circle model (Hamill & Gilbert 2009), which is easily incorporated into an ABM framework (zu Erbach-Schoenberg et al. 2013).

The Influence, Susceptibility, and Conformity Model (ISC)

- 3.1** To summarize, agents are randomly placed within a two-dimensional space. Each agent has a unique initial opinion, three parameters for tolerance, conformity, and susceptibility, and a social network. Each round, every agent initiates a dialogue with members of his social network. In the dialogue, each agent expresses an opinion that reflects his true opinion, his conformity, and the opinions already expressed in the dialogue. Afterwards, the initiating agent updates his true opinion based on his tolerance, susceptibility, and the expressed opinions' weighted influences. The model records the true and expressed opinion of each agent after every round. The code is written in Python 2.7; the model, data, and figures are available on GitHub.
- 3.2** Agents' opinions, interpreted as beliefs on a single subjective issue, lie on a continuous 0 – 100 scale. Initial opinion, tolerance, conformity, susceptibility, and social reach are all drawn from normal distributions whose means and variances are specified in each experiment. Agents are randomly assigned a continuously-valued (x, y) location, then each agent creates a social network N with all agents within euclidean radius equal to his social reach r , as per the social circle model. Agents remain stationary.
- 3.3** Agent i initiates a dialogue with all agents j in his social network. He is the first to express an opinion, and always voices his true opinion (O_i). Subsequently, each j distorts his opinion in order to conform or appear distinct. Specifically, j calculates the average of all opinions (E_k) expressed so far in the dialogue (D), then expresses an opinion (E_j) that is between his true opinion (O_j) and the dialogue's opinion norm (conformity), or that is distanced by some amount from the dialogue's norm (distinctiveness):

$$E_j = O_j + \frac{c_j}{k_j} * \frac{1}{N} \sum_k^D (E_k - O_j) \quad (1)$$

The agent parameter c_j represents an agent's inherent willingness to misrepresent his beliefs in social contexts in order to appear either normal or distinct. The parameter captures both conformity ($c_j > 0$) and distinctiveness ($c_j < 0$). Greater magnitude c_j produces greater belief falsification: $c_j = 0$ causes the agent to speak truthfully, $c_j = 1$ causes the agent to express the dialogue's "mean opinion", and $c = -1$ causes the agent to express an opinion that is more dislike the mean than his true opinion. In this model, conformity and distinctiveness are manifest in expression but not directly in opinion change: agents attempt to gain social favor by stating opinions that differ from their true beliefs, but do not change their true beliefs to reflect this posturing. This conceptually differentiates peoples' internal beliefs from external indicators of their opinions, which helps explain phenomenon such as groupthink (Janis 1972) in which public conformity belies undercurrents of dissent, suppresses original thinking, and leads to sub-optimal outcomes. This formulation of social context also differs significantly from previous models of conformity (Smaldino & Epstein 2015) and distinctiveness (Mäs et al. 2014), in which the opinion norms of the population directly influence ΔO_i .

- 3.4** The extent of j 's conformity is further mitigated by his current commitment k_j , which is proportional to his susceptibility s_j and the extremeness of his current opinion:

$$k_j = 1 + s_j * \frac{|50 - O_j|}{50}. \quad (2)$$

The susceptibility parameter s_j represents an agent's inherent commitment to strong beliefs; it causes him to be less affected by social context and social influence. Its magnitude governs how a departure from a neutral opinion ($O_i = 50$) translates to a shrinking of influence: higher values result in less opinion change. Opinion-dependent susceptibility gives some conceptual significance to the ideological spectrum by making the absolute value of an opinion meaningful: whereas previous models assumed that agent behavior only depended on opinion differences, this mechanism makes agents with extreme opinions behave differently than those with centrist opinions.

- 3.5** After each j has expressed E_j once in the dialogue, i updates his true opinion according to the dialogue's influence (I_i), which is proportional to each E_j and the weight that i assigns to that expression (w_{ij}):

$$I_i = \frac{\sum_j^N w_{ij} * (E_j - O_i)}{\sum_j^N |w_{ij}|} \quad (3)$$

Conceptually, the dialogue's influence I_i results from i being pulled towards (or pushed away from) each opinion expressed in the dialogue, E_j , by an amount proportional to the interagent weight, w_{ij} . The weight, in turn,

is calculated according to homophily: the greater the absolute distance between i 's opinion and j 's expression, the more negative the weight, and the less influence j 's expression will exert on i 's opinion:

$$w_{ij} = 1 - t_i \frac{|E_j - O_i|}{50} \quad (4)$$

where t_i represents i 's inherent intolerance of dissimilar opinions. Its magnitude dictates how strongly a given opinion difference translates to a loss of interagent weight. A high value implies that an agent will only assign positive weight to opinions that are similar to his own beliefs; a low value implies the agent will be positively influenced by a wider range of opinions. Mathematically, t_i is the slope of i 's weight vs. Δ opinion curve, which is continuous and linear. This is a departure from the canonical bounded confidence approach, in which weight is a threshold function of an agent's intolerance ϵ_i (Hegselmann et al. 2002). I believe continuous weighting better reflects the subtleties of opinion appraisal and social influence than a binary "full acceptance vs. complete disregard" judgment, while still capturing the essential idea of homophilous influence. This approach has also been adopted by (Mäs et al. 2014). Weights are bounded from -1 to $+1$.

- 3.6** Finally, i updates his true opinion based on his previous opinion and the dialogue's influence, scaled by his commitment:

$$O_{i,t+1} = O_{i,t} + \frac{I_i}{k_i} \quad (5)$$

This process is repeated for each i in the population, concluding one timestep.

- 3.7** I use four metrics to investigate the diversity, dynamics, and clustering of opinions within the population. Opinion histograms plot the frequency of opinions across the ideological spectrum at particular times, and are the most complete measure of strong vs. weak diversity. Opinion trajectories plot each agent's history as a line on a opinion vs. time graph, and are used to study dynamics towards or away from diversity. To distinguish different regions of opinion space, I use the terms **centrist** to describe agents who hold ($33 < O_i < 66$), **moderate** to describe agents with moderately-strong opinions ($16 < O_i < 33$ or $66 < O_i < 83$), and **extremist** to describe agents with the strongest opinions ($0 < O_i < 16$ or $83 < O_i < 100$). Spatial maps plot each agent as a circle at (x, y) in the model's abstract "network space" with color representing the agent's opinion, and can help identify subgroup formation and regions of ideological mixing. Distance in network space represents the degree of separation between two agents, which can be interpreted geographically or digitally depending on the communication medium. Finally, the **Jensen-Shannon Divergence** (JSD) is a measure of the similarity of two opinion distributions. I use the JSD to (a) quantify pluralistic ignorance and study how agents' falsifications affect the diversity of opinions within the population, and (b) compare the similarity of model and empirical opinion distributions. The JSD is calculated from the Kullback–Leibler divergence $D(P||Q)$, a standard entropy metric for probability distributions:

$$JSD(P||Q) = \frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M) \quad (6)$$

where P and Q are the true and expressed opinion distributions, $M = \frac{1}{2}(P + Q)$, and

$$D(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}. \quad (7)$$

The JSD ranges from 0 (identity) to 1 (minimum mutual information).

- 3.8** In the computational experiments below, I investigate single realizations of the model as artificial case studies rather than investigating the average behavior over many realizations. This is necessary to determine cause and effect in the simulation given highly-variable initial conditions and nonlinear interactions between agents. It is also a reasonable approach given that my goal is to demonstrate the emergence of strong diversity, opinion subcultures, and pluralistic ignorance in the model, rather than characterize the model's behavior across the parameter space.

Results: Increases in Psychological Complexity

- 4.1** In this section, I begin with the simplest version of the model and show that each additional psychological force expands the range of possible dynamics. This situates the model with reference to a lineage of existing models (while reviewing their results and limitations) and justifies its increased complexity. Readers familiar with the opinion dynamics literature may safely skip this section and proceed directly to an exploration of the model's

novel behaviors. For these preliminary experiments I assume no heterogeneity of intolerance, susceptibility, or conformity ($\sigma_t, \sigma_s, \sigma_c = 0$) and a social reach of $\mu_r = 22, \sigma_r = 4$ (for a 316×316 map).

Social influence and intolerance

- 4.2** I first reproduce a classical experiment in opinion dynamics, in which the final distribution of opinions is examined as a function of intolerance. In this experiment, other psychological forces are absent, with the exception of distancing, which is inherent in Equation 4. The result from numerous bounded confidence models is that low intolerance promotes societal opinion convergence, while high intolerance produces opinion polarization and weak diversity.
- 4.3** In this model, intolerance is an agent-level parameter t_i which is initially drawn from a normal distribution of mean μ_t and variance σ_t . In a population with low intolerance, $\mu_t = 0.7$, most agents assign positive weight to each others' opinions during dialogues, and are consequently pulled towards the mean opinion in that dialogue. Figure 1 (left) shows that an initial normal distribution of opinions, $\mu_O = 50, \sigma_O = 20$, rapidly converges to a single, centrist opinion: given enough time, diversity will completely disappear, and all agents will believe $O_i = 50$. Conversely, in a population with high intolerance, $\mu_t = 1.0$, many agents assign negative weight to each others' opinions and are pushed away from the dialogue mean. As agents adopt stronger opinions, they assign stronger negative weights, resulting in polarizing feedback. Figure 1 (right) shows this population rapidly diverges to two extremists opinions at either end of the opinion spectrum. As $t \rightarrow \infty$ only weak diversity remains: all agents either hold $O_i = 0$ or $O_i = 100$. These base-case results confirm the classical finding that, in the absence of other psychological forces, the degree of individuals' intolerance determines whether the population homogenizes or polarizes. Additional runs show that societies with intolerance below $\mu_t = 0.7$ always converge, societies with intolerance above $\mu_t = 1.0$ always diverge, and societies in between can either converge or diverge, depending on initial conditions. Notably, strong diversity is never obtained, as in bounded confidence models.

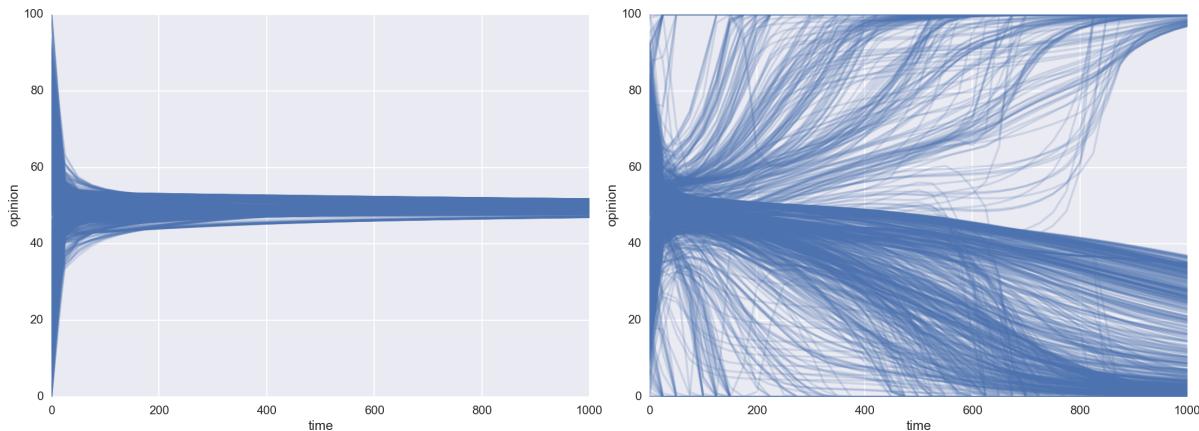


Figure 1: A population agents who initially hold normally distributed opinions converges to a single centrist opinion if agents' intolerance of dissimilar opinions is low, and diverges to two extreme opinions if agents' intolerance is high. This result reproduces findings in classical opinion dynamics and represents the base case of the simulation, in which social influence is the only active psychological force and all agents are identically intolerant.

Conformity and distinctiveness

- 4.4** Next, I introduce social context into the simulation by allowing agents to misrepresent their true opinions in dialogues. This mechanism extends previous studies by (a) investigating the interaction of distancing and conformity/distinctiveness, which were previously explored as stand-alone forces to oppose positive, homophilous social influence (Mäs et al. 2014), and (b) stipulating that social context does not directly affect agents' true opinions, but instead causes them to verbally falsify their opinions, which affects the information available to other agents. I hypothesize that if falsification is significant, agents will perceive an unrepresentative distribution of opinions and change their beliefs accordingly: a conformist population will homogenize under conditions that otherwise cause polarization, while a population driven by distinctiveness will polarize under conditions that otherwise favor consensus.

- 4.5** First, I simulate a population whose high intolerance would normally cause polarization, $\mu_t = 1.0$, but introduce a moderate tendency towards social conformity, $\mu_c = 0.5$. Conforming agents now express opinions that are close enough to the dialogue mean other agents never assign these (falsified) opinions a negative weight. Agents adopt opinions closer to the norms expressed in the dialogue, and opinions converge, eventually resulting in societal consensus as shown in Figure 2 (right). Second, I simulate the opposite conditions: a population whose low intolerance would normally cause convergence, $\mu_t = 0.7$, but filled with agents who possess a strong desire to be distinct, $\mu_c = -1.5$. When agents converse, they express opinions that differ radically from centrist norms. As expressions become more extreme, social influence causes agents to adopt and retain extreme beliefs. Figure 2 (left) shows that opinions initially converge, then diverge towards two homogeneous extremist parties. These results indicate that contextual opinion falsification can reverse the effects of intolerance, but cannot sustain strong diversity or pluralistic ignorance.

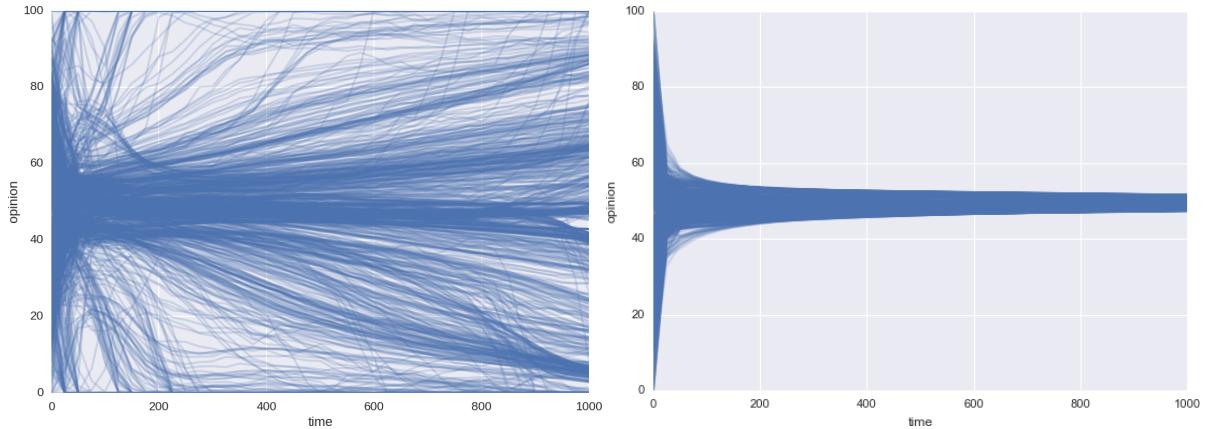


Figure 2: When agents with low intolerance wish to appear distinct in social contexts, they express extremists beliefs in dialogues, which eventually polarizes the population and leaves weak, bimodal diversity. When agents with high intolerance are motivated to socially conform, they express normative centrist views, causing the population to converge to centrism.

Commitment to strong beliefs

- 4.6** I conclude the preliminary experiments by investigating whether, when agents' susceptibility to influence decreases with their belief extremity, different patterns of societal opinion change emerge. I hypothesize that when extremist agents undergo less opinion change than moderate or centrist agents, their persistent influence will prevent centrist homogenization and polarize society.
- 4.7** Beginning with the simple case of a tolerant population with no opinion falsification, I test whether strong commitment, $\mu_s = 10.0$, can reverse trends towards convergence. The mean opinion expressed in dialogues is still $\bar{O} \approx 50$, but because agents are tolerant and truthful, they assign positive weights to all opinions they hear. Extreme agents undergo little opinion change due to their commitment, but without a repelling force to push them away from social norms (i.e. intolerance or distinctiveness), they are still pulled slowly towards this centrist opinion. Although they remain steadfast in their views for longer periods of time than in the “social influence and intolerance” experiment, they eventually converge to a single centrist opinion like the rest of the population (not shown). This result shows that commitment by itself cannot reverse homophilous opinion convergence; a distancing mechanism is still required.
- 4.8** However, personal susceptibility can affect a population that is intolerant and conformist, which normally homogenizes as in the first case of the “conformity and distinctiveness” experiment. Opinions initially converge due to the strong centrist norms perceived in conformist social dialogues, but extremists are slow to change. By $t = 300$, most agents have adopted moderate or centrist opinions and expressions, but about 2% of agents have, through intolerant repelling, adopted maximally extreme opinions, as shown in Figure 3. These extreme agents are now so committed to their beliefs that they barely soften their expressions to socially conform, $O_i = 100 \rightarrow E_i = 95$, and their strongly opinionated vocalizations polarize their social networks. Over time, this influence bifurcates the population, as can be seen by the divergence of opinions past $t = 500$. This experiment indicates that personal commitment fosters pockets of extremism whose long-term influence significantly alters the population's opinion dynamics. Although previous models have shown that extremists can spread polarization through social networks and instigate large-scale change (Jarman et al. 2015; Deffuant et al. 2002),

such models often initialize extremist agents with special properties or rules: here, extremists emerge (and disappear) endogenously from a population governed by universal rules but initialized with heterogeneous attributes.

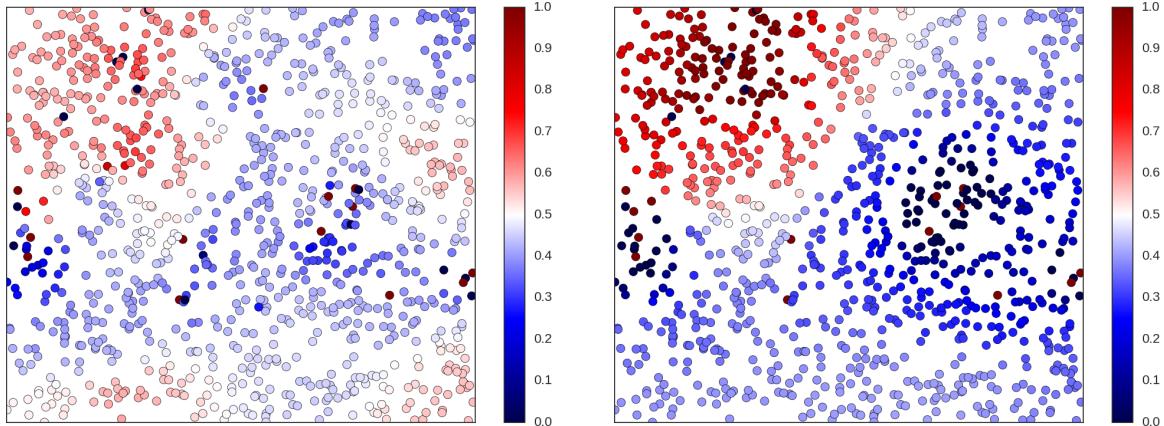


Figure 3: In a population filled with individuals who are intolerant, conformist, and committed to extreme beliefs, a minority of agents will distance themselves from centrist norms and stubbornly express extremist views. The influence of these agents polarizes their neighbors, spreading extremism spatially outward until the population bifurcates into weak diversity. Spatial maps at $t = 200$ and $t = 2000$ (size 316×316) show that polarization originates from neighborhoods that contain extremist agents.

Results: Novel Behaviors

- 5.1** In this section, I simulate populations in which all three psychosocial forces interact. In the following experiments, all agent parameters are drawn from normal distributions with nonzero means and variances, creating a population with greater heterogeneity and psychological realism than previous opinion dynamics models. I hypothesize that when agents are simultaneously motivated by social influence, personal susceptibility, and social context of varying degrees, the population will (a) maintain a strong diversity of opinions, (b) self-organize into opinion subcultures, and (c) exhibit pluralistic ignorance.

Strong diversity

- 5.2** First I examine a population which is on average tolerant and conformist ($\mu_t = 0.7, \mu_c = 0.4$), but with enough diversity to create some intolerant and distinction-seeking agents ($\sigma_t = 0.3, \sigma_c = 0.1$). These forces tend to pull population towards convergence. However, an opposing commitment to strong beliefs ($\mu_s = 5.0, \sigma_s = 0.4$) helps initially-extreme agents resist normalization and locally exert polarizing influence. The opinion trajectory plot in Figure 4 shows that the population initially converges, but scattered extremists retain their strong views. Unlike in the previous experiment, where strong commitment kept extremists from softening their expressions, extremists in this population moderate their expressions and few extreme opinions are publicly voiced. This, combined with low extremist density, prevents radicals from attracting many followers: by $t = 1,000$, less than 5% of the population holds or expresses extreme views. However, these extremists exert enough influence that they keep centrist norms from completely homogenizing the population. By $t = 10,000$, the distribution of opinions and expressions have settled into a diverse centrist group, a moderate fringe, and scattered extremists. This strongly diverse distribution of opinions persists past $t = 100,000$ despite small opinion fluctuations. A spatial map of opinions shows that the diversity within the centrist party arises from the minor influence exerted by extremists, which keeps the surrounding neighborhoods to the ideological left and right of $\bar{O} \approx 50$.

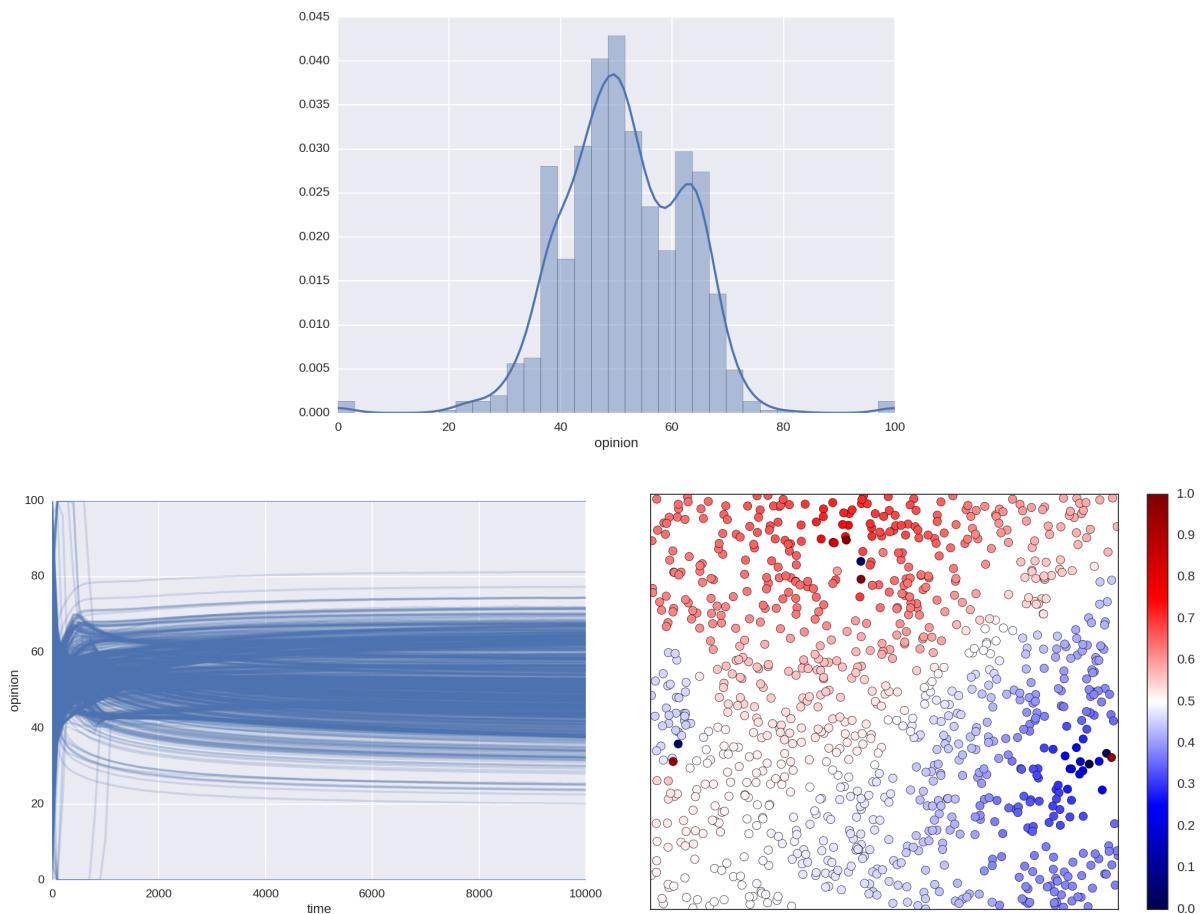


Figure 4: When agents are motivated by social influence, personal susceptibility, and social context of heterogeneous strength, novel opinion distributions emerge at the societal scale. The opinion trajectory shows that the population settles into a stable opinion configuration, while the opinion histogram and spatial map confirm that the final distribution of opinions is strongly diverse. Most agents have adopted a centrist opinion, but a small minority of extremists counterbalance homogenizing norms, preventing total convergence but exerting too little influence to bifurcate the population

5.3 This result is, to my knowledge, the first evidence of indefinitely-sustained strong diversity in a continuous-opinion model. When intolerance, conformity, and susceptibility simultaneously motivate agents' behavior with heterogeneous strength ($\sigma_{t,c,s} > 0$), agents experience a multitude of forces simultaneously pushing their opinions and expressions in many directions. When homogenizing and polarizing forces balance across an entire population, agents settle into a relatively stable opinion configuration in which individual opinions and expressions still fluctuate, but not in a unified direction. Interestingly, this delicate balance only occurs when all three forces are present (with heterogeneous strength): simpler versions of the model presented in the previous section always ended with convergence or weak diversity, as did other unreported simulations which varied $\mu_{t,c,s}$ over a wide range. This supports the central hypothesis that strong diversity emerges from the increased psychological complexity of the model.

Opinion subcultures

5.4 Next, I increase agents' average intolerance and conformity ($\mu_t = 1.0, \mu_c = 0.5, \mu_s = 5.0$), then tweak their psychological diversity ($\sigma_t = 0.3, \sigma_s = 0.3, \sigma_c = 0.3$). The results are shown in Figure 5. After an initial period of convergence, several extremists neighborhoods develop, affecting partial polarization that quickly spreads outwards. Eventually, the population settles into a small centrist group and two dominant extremist groups that continue to influence each other through persuadable agents on their mutual border. The spatial orientation of the extremist parties is such that the centrist party receives approximately equal influence from both sides of the opinion spectrum, and acts as a relatively stable buffer between the two extremes.

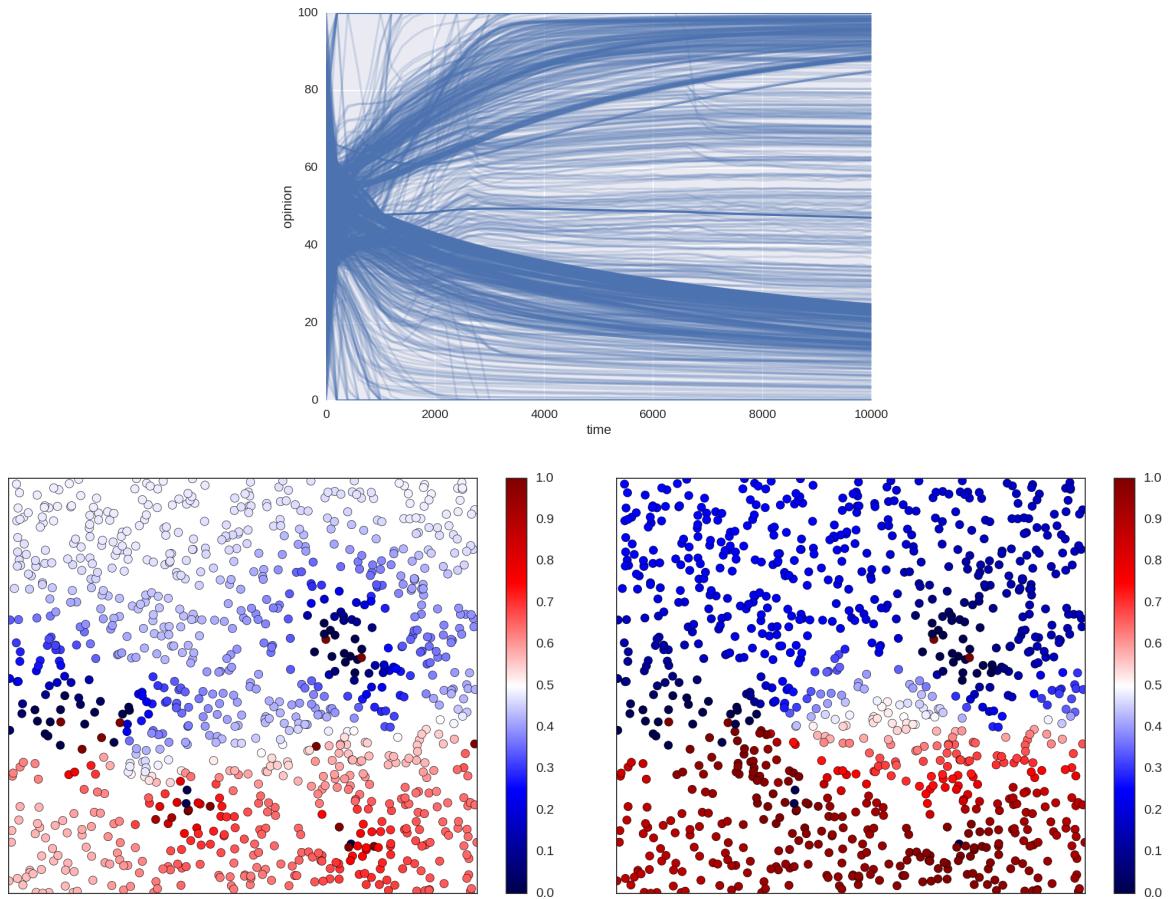


Figure 5: Extremists counteract initial trends towards convergence and form neighborhoods of strongly (but uniquely) opinionated agents. These groups compete on the border: when one group exerts greater influence, they persuade moderate agents to become extremists; and when both groups exert equal influence, a buffer zone of centrist forms between them. These strongly diverse subcultures persist through time.

- 5.5** Opinion subcultures should also depend on the size of agents' social networks. I hypothesize that small social networks will promote distinct opinion subcultures while large social networks will dissolve subgroups. I reran the above experiment with smaller social networks, obtained by reducing agents' social reach (from $\mu_r = 22$, $\sigma_r = 4$ to $\mu_r = 11$, $\sigma_r = 2$). Opinions rapidly cluster according to social networks, producing discrete subcultures, Figure 6. These subcultures continue to receive influence from surrounding networks, often through persuadable agents on the border who continually oppose consensus and promote strong diversity within each subculture. The partial isolation of subcultures prevents both homogenization and polarization globally, which is reflected in a wide, multimodal opinion distribution. Both opinion subcultures and strong diversity remain stable past $t = 5,000$.

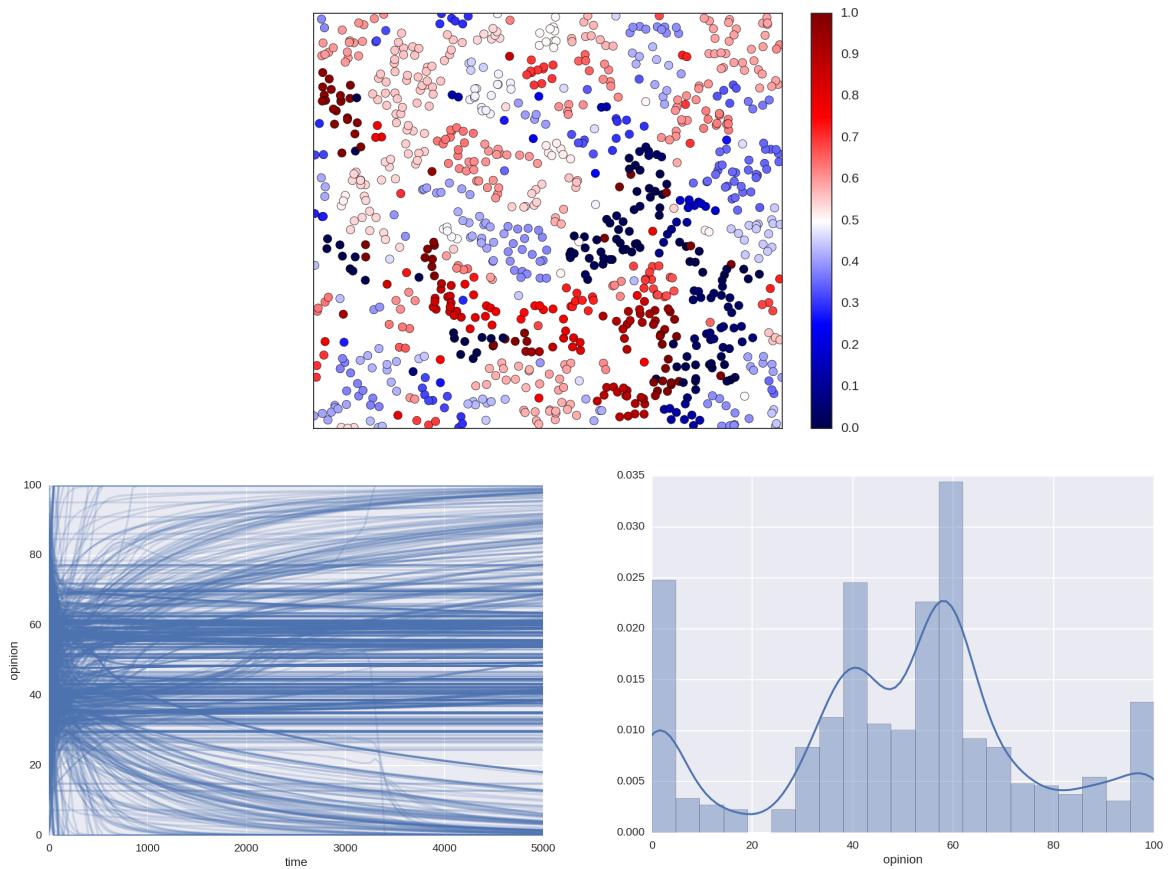


Figure 6: Shrinking agents' social networks encourages the formation of clustered opinion subcultures. Reduced interaction between these groups prevents both centrist and extremist takeover, but continuing dialogues with intermediary agents keeps these groups, and the population as a whole, strongly diverse. Histograms and maps show opinions at $t = 10,000$.

5.6 Conversely, increasing agents' social reach promotes homogenization. Initially, most agents converge to centrism, while the large size and strong centrist norms in dialogues encourage the remaining extremists to express moderate views. Two outcomes are possible: either all extremists convert and the population converges to centrism; or, as shown in Figure 7, an imbalance of extremists remains, and the population drifts towards the most vocal group. Unlike in the first “opinion subcultures” experiment, where a buffer zone between extremist groups prevented takeover, large networks decrease the likelihood that centrist agents participate in dialogues that are well-balanced between the two extremes. This increases the probability that the dominant group will exert the strongest influence across all social networks, and that centrists and moderates will turn towards that extreme. In either case, strong diversity vanishes.

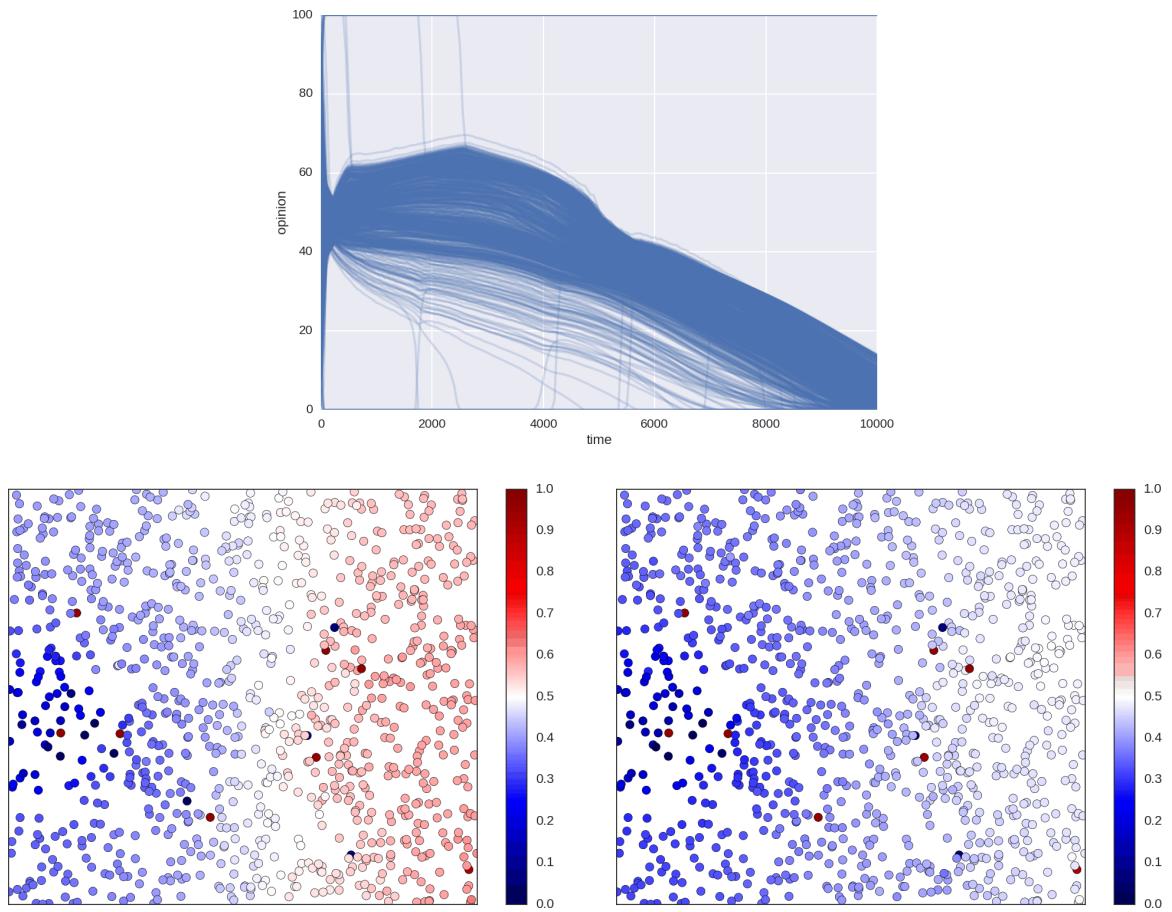


Figure 7: Expanding agents' social networks dilutes the influence of extremist agents over larger network, while globalized influence prevents the formation of opinion subcultures and eventually destroys strong diversity. The spatial maps at $t = 4,000$ and $t = 5,000$ show that the buffer zone that previously preserved diversity no longer prevents the takeover of the dominant extremist group.

5.7 Collectively, these results are interesting for two reasons. First, they show that subcultures can persist despite mutual interaction between dissimilar agents and through uncommitted agents that bridge the ideological gap. This can be contrasted with subcultures in bounded confidence models, which typically survive because agents become too dissimilar to influence one another at all, disconnecting the influence graph. Second, these simulations have the same average intolerance, conformity, and susceptibility as the “commitment to strong beliefs” experiment, but unlike that experiment maintained strong diversity or drifted to one extreme rather than bifurcating. This demonstrates that psychological diversity (heterogeneous intolerance, conformity, susceptibility) and the size of social networks (social reach) within a population can substantially affect large-scale opinion dynamics.

Pluralistic ignorance

5.8 Pluralistic ignorance and unpredictable dynamics are also possible under various conditions, such as when agents have intermediate intolerance ($\mu_t = 0.8, \sigma_t = 0.3$), low commitment ($\mu_s = 0.1, \sigma_s = 0.1$), and highly variable conformity ($\mu_c = 0.3, \sigma_c = 0.5$). Opinions converge early on, and the population is sufficiently tolerant and uncommitted that only a few agents retain extreme opinions. Through some combination of the extremists' social influence, conformity of their neighbors, and distinction of agents from centrist norms, opinions throughout the population begin drifting towards the extreme. However, the extremists abruptly convert to centrism, causing a dramatic turn towards convergence, Figure 8. Before these conversions, centrist and extremist agents falsely express moderate opinions in dialogues, leading to high pluralistic ignorance. The perceived moderate norm pulls centrists towards extremism, causing a slow drift towards extremism before $t = 500$, but also pulls extremists towards centrism, causing the occasional conversion of an extremist agent. If the former trend dominates, the population bifurcates; if the latter dominates, the population homogenizes.

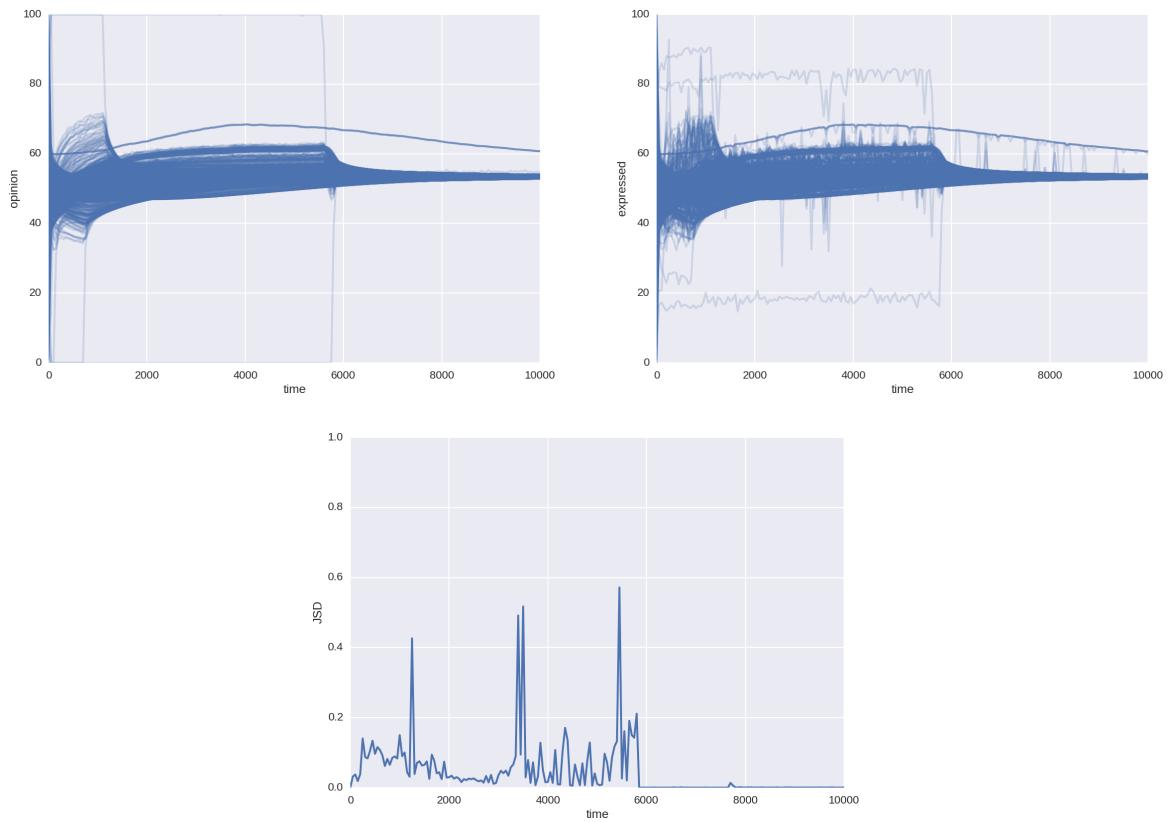


Figure 8: The conversion of influential extremists may shift the course of societal opinion change from polarization to convergence. These events occur rapidly and are frequently preceded by spikes in pluralistic ignorance, as demonstrated by a large Jensen-Shannon Divergence (JSD) between the true and expressed opinion distribution at $t \simeq 500, 1000, 5800$. This suggests that when sustained levels of opinion falsification are finally revealed, tipping-point phenomenon may occur, leading to nonlinear opinion dynamics.

- 5.9** This result suggests that high pluralistic ignorance may precede dramatic changes in societal opinion dynamics, a conclusion which aligns with empirical instances of unvoiced discontent preceding popular revolutions in authoritarian societies (Kuran 1989; Goodwin 2011). The notion that dramatic, macroscopic opinion shifts are instigated by the conversion of extremists (e.g. key party or military leaders defecting from the regime and supporting revolutionary groups) is further supported by the literature on authoritarian-democratic transitions (O'Donnell et al. 2013).

Validation: American Political Opinions

- 6.1** Although the ISC model is grounded in social psychology and reproduces features like strong diversity, opinion subcultures, and pluralistic ignorance, I have not shown that it quantitatively captures real-world data. In this section, I validate the model by reproducing empirical data on the distributions and dynamics of political opinions in American society.
- 6.2** As a proof-of-concept for strong diversity, I compare the expressed opinion distributions produced by the ISC model with a survey that assessed people's opinions on each of twelve issues in contemporary American politics (Broockman 2016). Each respondent was asked which of seven idealized positions, ranging from extremely liberal to extremely conservative statements about that issue, best described his or her belief, creating a seven-point opinion scale. Using several parameter-space exploration strategies, including an evolutionary algorithm and hyperopt (Bergstra et al. 2015), I found values for mean intolerance, susceptibility, and conformity that produced the distributions shown in Figure 9. These parameters, optimized to reduce the root-mean-square-error between the model distribution and Broockman's data over $n = 4$ realizations, lie within the bounds of the values used in the above experiments (except for $\sigma_O = 50$). The model captures distributions with a variety of different shapes, including: normal distributions around a centrist opinion (gun control) and a moderate opinion (affirmative action); centrist dominance with an extreme group (healthcare and contraception); and

other strange shapes (abortion and immigration, with less accuracy). This result quantitatively demonstrates that the strong diversity found in real-world opinion distributions are within the output-space of the model.

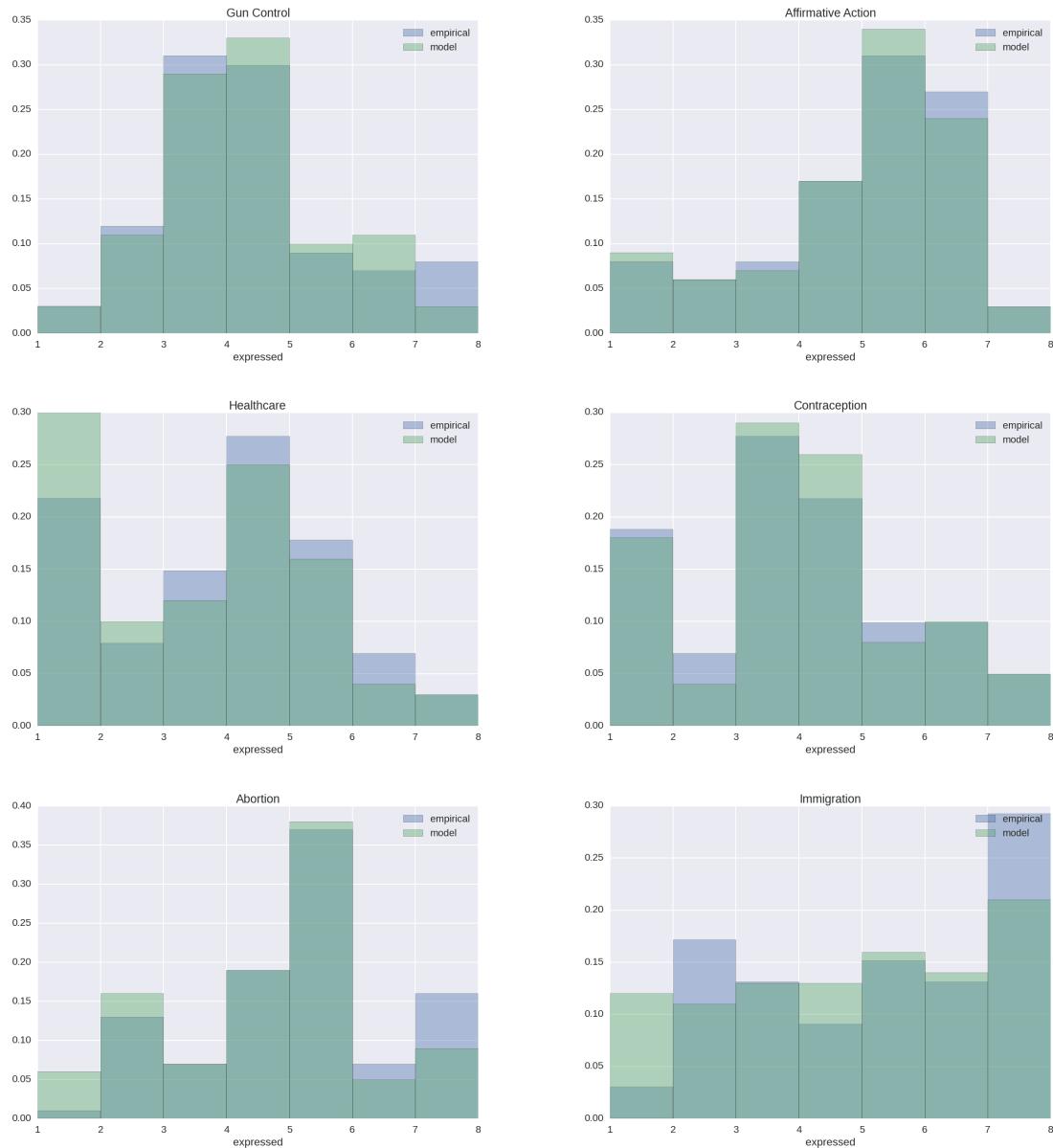


Figure 9: The ISC model produces expressed opinion distributions that align with American's opinions on contemporary political issues ranging from gun control to healthcare to immigration. The JSD between the model and empirical distributions ranges from 0.017 (affirmative action) to 0.038 (healthcare). Data reproduced with permission from Broockman (2016), parameters for each realization available on GitHub

6.3 I also compare the model's opinion dynamics with a large- N , multi-year survey of American's ideological consistency conducted by the Pew Research Center (Dimock et al. 2014). The survey consisted of ten questions assessing individuals' attitudes about current political issues such as "[the] size and scope of government, the social safety net, immigration, homosexuality, business, the environment, foreign policy and racial discrimination," with each response coded -1 (liberal), $+1$ (conservative), or 0 (don't know/refused). These values were summed for each individual, creating an "ideological consistency" scale ranging from -10 (liberal responses to every question) to $+10$ (conservative responses to every question). The study found that Americans have become increasingly polarized from 1994 to 2014: individuals who previously held mixed liberal and conservative positions on different issues are increasingly partisan and ideologically uniform. As shown in Figure 10, this trend manifests not as a bifurcation of opinions but as a spreading of the empirical distribution over time. Using the same parameter-space exploration tools as above, I found the ISC model produced similar patterns

of polarization: a normal-like opinion distribution measured midway through the simulation gradually spread as extremists on both sides pulled centrists towards the periphery. The model also reproduced more subtle dynamics, such as the leftward shift of the kernel density estimate's central peak from 1994 to 1999, then back to the right as a sharper peak from 1999 to 2004.

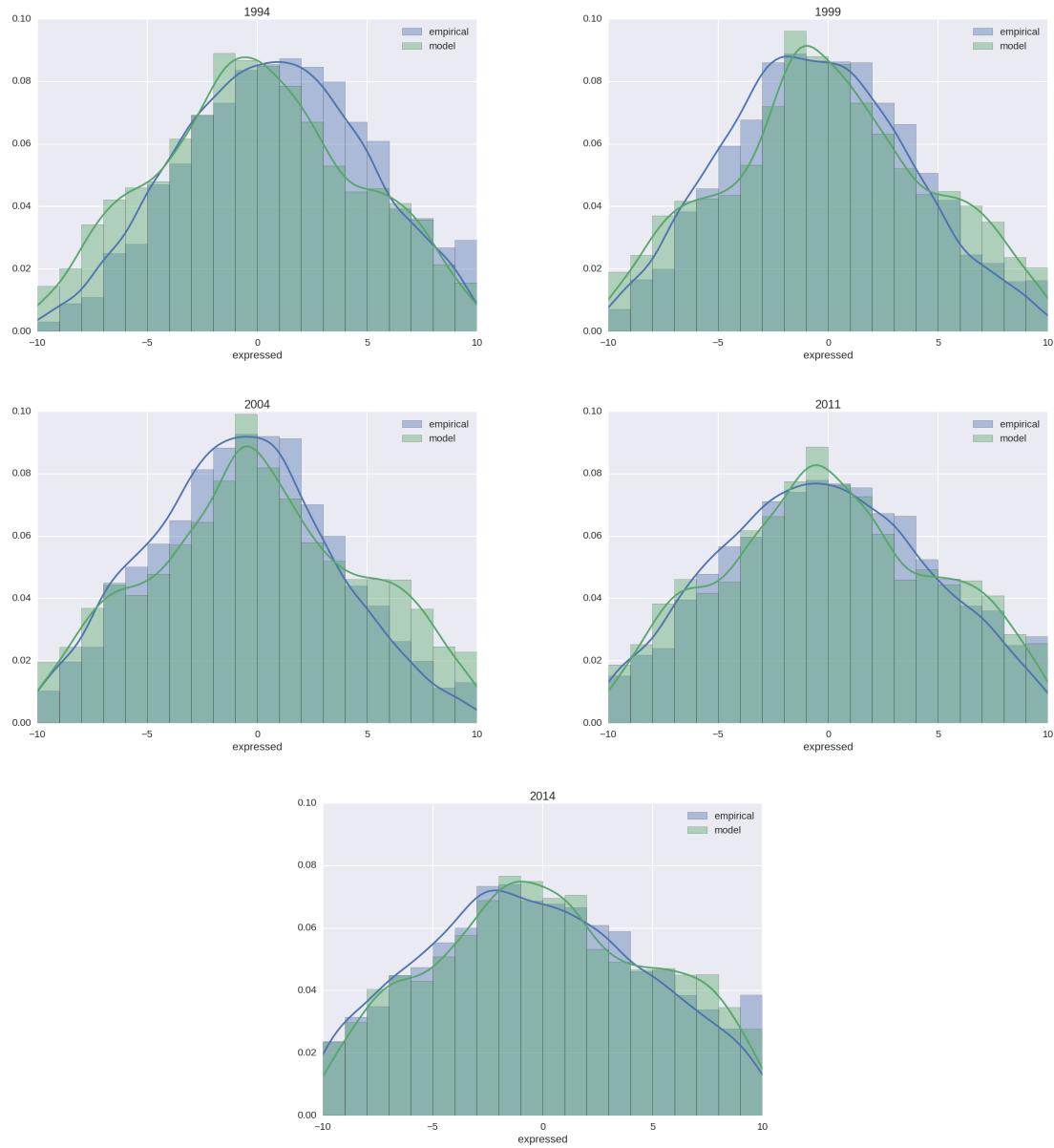


Figure 10: The ISC model produces opinion dynamics that are consistent with the polarization of Americans' political opinions from 1994 to 2014. Blue and green lines are Gaussian kernel density estimates for the respective distributions. The average JSD between the model and empirical distributions ranges over the five data points was 0.01. Data reproduced from the Pew Research Center Dimock et al. (2014).

- 6.4** All models should be treated with skepticism until they have been credibly validated with empirical data. Overall, these validation results should be seen as an existence proof of plausible diversity and dynamics in the model, not as evidence of a calibrated simulation capable of precisely predicting opinion change.

Discussion

Why do societies maintain a diversity of opinions?

- 7.1** Psychological forces such as a commitment to strong opinions or a drive to distinctiveness may oppose homogenizing social influence and preserve strong diversity. When the population contains agents with heterogeneous intolerance, susceptibility, and conformity, each individual is simultaneously pulled towards centrism and extremism. If these forces are balanced, which occurs in this study only when all three forces are present, a strong diversity of opinions emerges and remains stable through time. Figures 4, 5, and 6 showed this diversity can take the form of (a) a centrist party diversified by influence from a few extremist, (b) two extremist parties with undecided agents on the borders, and (c) opinion subcultures clustered according to social networks.
- 7.2** The maintenance of strong diversity is a novel result in opinion change and cultural diffusion models based on bounded confidence, which assume that agents influence one another only if their opinion similarity is above an interaction threshold (Macy et al. 2003; Hegselmann et al. 2002; Deffuant et al. 2002; Dandekar et al. 2013; Salzarulo 2006; Jager & Amblard 2005). Though this approximation of intolerance has provided insights into convergence and polarization, I have argued that it is overly rigid: people do not classify each others' trustworthiness according to a binary scheme. The ISC model assumes that social influence changes continuously with intolerance, commitment, and context, and produces sustained, strong diversity under multiple psychological and network conditions. This result is intuitive, since societies do not converge to a single opinion or diverge to two polar opposites, and is also quantitatively plausible, as shown through empirical validation. Whenever possible, agent-based modelers should move away from psychologically and socially implausible assumptions and adopt empirically-motivated cognitive heuristics: doing so will solidify the model's foundations and, as exemplified by this study, produce more complex and realistic results.
- 7.3** Maintaining a diversity of opinions is important outside the modeling community. Indigenous cultures dissolve in the face of globalization as people substitute traditional languages and practices for the norms of modern society (Arnett 2002). Corporations fall prey to groupthink when individuals with original ideas choose not to voice them (Janis 1972). Political and religious groups become polarized due to intolerance of dissimilar beliefs (Lord et al. 1979; Taber & Lodge 2006). To promote cultural and ideological diversity, leaders must recognize that social influence is not the only force that drives single-mindedness. They must recognize, not just conceptually but with the quantitative precision afforded by computational models, the role of psychological heterogeneity, personal commitment, and social context in destroying the valuable resource of diversity.

Will subcultures of opinions survive in a well-connected population?

- 7.4** The formation and survival of opinion subcultures, as well as the spread of extremism, depends critically on how influence spreads through social networks. In the ISC model, communities of dissenters can survive among globalized centrism or extremist competition in two circumstances. In the first scenario, intolerant agents strongly reject opposing beliefs, becoming extreme and spreading polarization outward. This leads to the formation of cohesive extremist parties with undecided agents on the neighborhoods' borders. Densely clustered extremists resist moderate influence: neither a centrist majority nor an opposing extreme minority effectively moderates their speech or prevents their polarizing influence. In the second scenario, agents' small social networks limit communication, producing a larger number of semi-isolated, cohesive, persistent communities. Communication is still possible between such communities, but must travel through bridging individuals whose influence is often overcome by the group consensus. These communities cannot coalesce when agents' social networks are large; social influence, when distributed over a large network, may cause either centrist convergence or extremist takeover.
- 7.5** A lack of communication can encourage ideological splintering in the same way that geographic barriers facilitate speciation and genetic diversification. For instance, extremism often emerges in online communities with limited access to external influence or information (Radlauer 2007). The ISC model provides a psychosocial explanation for this phenomenon; the model also suggests that, although networking extremists with moderate individuals has the potential to spread radicalization, it also provides a bridge to moderate attitudes which may gradually soften extremists' beliefs. Organizations seeking to prevent the spread of religious, political, or ethnic extremism (Ressler 2006) would benefit from examining radical recruitment and deradicalization strategies through the lens of a psychologically-motivated opinion change models.

How do hidden opinions affect societal opinion change?

- 7.6** Social norms like the desire to be distinct from previous generations or to conform to a community's religious beliefs have far reaching effects on opinion change at the societal level. The ISC models shows that agents' conformity can lead others to mistakenly think agreement exists in the population, reverse the process of opinion polarization, and bring an intolerant population back to consensus. It also shows that when agents express opinions with the goal of appearing distinct, a false atmosphere of extremism can incite societal polarization.
- 7.7** Imperfect information is another important aspect of social communication, and the ISC model is the first of its kind to explicitly differentiate internal belief revision from socially-motivated belief communication. The model indicates that pluralistic ignorance precedes critical periods of opinion change, such as when stubborn extremist convert to centrism or when centrist express more extreme sentiments. One interpretation of these results is that opinion dynamics are relatively predictable when everyone communicates perfectly, but when social context encourages belief falsification, tensions between what is heard and what is felt build until they are unexpectedly and abruptly released. This interpretation is consistent with studies on the role of preference falsification in authoritarian revolutions (Kuran 1989), and was likely a contributing factor in the buildup to the Arab Spring (Goodwin 2011).

Conclusion

- 8.1** In this study, I examined the relationship between the psychosocial forces driving opinion change and the resulting distributions, dynamics, and clustering of opinions across a population of agents. This research extends previous studies in computational opinion dynamics by expanding the social and psychological depth of agents. Through a series of simulations, I showed that networks of heterogeneous agents will interact to produce (a) opinion distributions and dynamics that match political opinion data, (b) opinion subcultures perpetuated by extremists and conformity, and (c) tipping points induced by pluralistic ignorance. These results are significant advances in the study of macroscopic opinion change and suggest that modest increases in the complexity of agent models can produce opinion dynamics that align better with reality.
- 8.2** Many extensions of the ISC model are possible. People actively promote their opinions at rallies or online, while others join organizations that enforce their beliefs through coercion and punishment. Introducing social mechanisms for these behaviors would permit the study of collective action problems and suggest more specific strategies that leaders could take to achieve desired patterns of opinion change. Another extension would allow for dynamic social networks. Though the social reach procedure captures important statistics of social networks, the people with whom we converse change constantly. Introducing dynamic networking, possibly in an expanded virtual environment, would permit a more complete study of how opinions change in a society dominated by social media. I would also like to compare opinion geography and pluralistic ignorance to empirical data.
- 8.3** I contend that empirically-accurate patterns of opinion change only emerge when agents act according to plausible rules, and that modelers must expand the depth of agents' social cognition to explain complex social phenomenon. This is best achieved by endowing agents with human-like cognitive architectures capable of affecting perception, memory, emotion, attention, and communication. Several opinion change models have already incorporated neurally-inspired mechanisms to great effect (Schröder & Thagard 2013; Wolf et al. 2015). Recent advances in neural engineering suggest that building agents with artificial brains may soon be possible (Eliasmith et al. 2012). In future work, I plan to incorporate such artificial intelligences into social simulations.

Acknowledgments

I would like to thank Joshua M. Epstein and my colleagues at the Johns Hopkins Center for Advanced Modeling in the Social Sciences for their mentorship and support of this research. I also acknowledge the use of several indispensable software packages, including hyperopt (Bergstra et al. 2015), seaborn (Waskom et al. 2015), pandas (McKinney 2010), and pathos (McKerns et al. 2012).

References

- Abelson, R. P. (1964). Mathematical models of the distribution of attitudes under controversy. *Contributions to Mathematical Psychology*, 14, 1–160
- Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology*, 52(1), 27–58
- Amblard, F. & Deffuant, G. (2004). The role of network topology on extremism propagation with the relative agreement opinion dynamics. *Physica A: Statistical Mechanics and its Applications*, 343, 725–738
- Arnett, J. J. (2002). The psychology of globalization. *American Psychologist*, 57(10), 774
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. *Groups, Leadership, and Men*, (pp. 222–236)
- Axelrod, R. (1997). The dissemination of culture a model with local convergence and global polarization. *Journal of Conflict Resolution*, 41(2), 203–226
- Baldassarri, D. & Bearman, P. (2007). Dynamics of political polarization. *American Sociological Review*, 72(5), 784–811
- Bednar, J., Bramson, A., Jones-Rooy, A. & Page, S. (2010). Emergent cultural signatures and persistent diversity: A model of conformity and consistency. *Rationality and Society*, 22(4), 407–444
- Bergstra, J., Komer, B., Eliasmith, C., Yamins, D. & Cox, D. D. (2015). Hyperopt: A Python library for model selection and hyperparameter optimization. *Computational Science & Discovery*, 8(1), 014008
- Broockman, D. E. (2016). Approaches to studying policy representation. *Legislative Studies Quarterly*, 41(1), 181–215
- Carley, K. (1991). A theory of group stability. *American Sociological Review*, (pp. 331–354)
- Carroll, G. R. & Harrison, J. R. (2002). Come together? The organizational dynamics of post-merger cultural integration. *Simulation Modelling Practice and Theory*, 10(5), 349–368
- Centola, D., Willer, R. & Macy, M. (2005). The emperor's dilemma: A computational model of self-enforcing norms. *American Journal of Sociology*, 110(4), 1009–1040
- Christakis, N. A. & Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, 358(21), 2249–2258
- Cialdini, R. B. & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55, 591–621
- Dandekar, P., Goel, A. & Lee, D. T. (2013). Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110(15), 5791–5796
- De Sanctis, L. & Galla, T. (2009). Effects of noise and confidence thresholds in nominal and metric Axelrod dynamics of social influence. *Physical Review E*, 79(4), 046108
- Deffuant, G., Amblard, F., Weisbuch, G. & Faure, T. (2002). How can extremism prevail? A study based on the relative agreement interaction model. *Journal of Artificial Societies and Social Simulation*, 5(4), 1
- Dimock, M., Kiley, J., Keeter, S. & Doherty, C. (2014). Political polarization in the american public. Tech. rep., Pew Research Center
- Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y. & Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science*, 338(6111), 1202–1205
- Epstein, J. M. (2006). *Generative social science: Studies in agent-based computational modeling*. Princeton University Press
- Epstein, J. M. (2014). *Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science*. Princeton University Press
- Flache, A. & Macy, M. W. (2011). Local convergence and global diversity from interpersonal to social influence. *Journal of Conflict Resolution*, 55(6), 970–995

- Friedkin, N. E. (1984). Structural cohesion and equivalence explanations of social homogeneity. *Sociological Methods & Research*, 12(3), 235–261
- Friedkin, N. E. (1993). Structural bases of interpersonal influence in groups: A longitudinal case study. *American Sociological Review*, (pp. 861–872)
- Gawronski, B. & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological bulletin*, 132(5), 692
- Goodwin, J. (2011). Why we were surprised (again) by the Arab Spring. *Swiss Political Science Review*, 17(4), 452–456
- Hamill, L. & Gilbert, N. (2009). Social circles: A simple structure for agent-based social network models. *Journal of Artificial Societies and Social Simulation*, 12(2), 3
- Hegselmann, R., Krause, U. et al. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2
- Imhoff, R. & Erb, H.-P. (2009). What motivates nonconformity? Uniqueness seeking blocks majority influence. *Personality and Social Psychology Bulletin*, 35(3), 309–320
- Jager, W. & Amblard, F. (2005). Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change. *Computational & Mathematical Organization Theory*, 10(4), 295–303
- Janis, I. L. (1972). *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes*. Houghton Mifflin
- Jarman, M., Nowak, A., Borkowski, W., Serfass, D., Wong, A. & Vallacher, R. (2015). The critical few: Anticonformists at the crossroads of minority opinion survival and collapse. *Journal of Artificial Societies and Social Simulation*, 18(1), 6
- Kitts, J. A. (2006). Social influence and the emergence of norms amid ties of amity and enmity. *Simulation Modelling Practice and Theory*, 14(4), 407–422
- Kitts, J. A., Macy, M. W. & Flache, A. (1999). Structural learning: Attraction and conformity in task-oriented groups. *Computational & Mathematical Organization Theory*, 5(2), 129–145
- Kitts, J. A. & Trowbridge, P. T. (2007). Shape up or ship out: social networks, turnover, and organizational culture. *Computational and Mathematical Organization Theory*, 13(4), 333–353
- Klemm, K., Eguíluz, V. M., Toral, R. & San Miguel, M. (2003). Global culture: A noise-induced transition in finite systems. *Physical Review E*, 67(4), 045101
- Kuran, T. (1989). Sparks and prairie fires: A theory of unanticipated political revolution. *Public choice*, 61(1), 41–74
- Kuran, T. (1997). *Private truths, public lies: The social consequences of preference falsification*. Harvard University Press
- Lapinski, M. K. & Rimal, R. N. (2005). An explication of social norms. *Communication Theory*, 15(2), 127–147
- Lord, C. G., Ross, L. & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology*, 37(11), 2098
- Macy, M. W., Kitts, J. A., Flache, A. & Benard, S. (2003). Polarization in dynamic networks: A Hopfield model of emergent structure. *Dynamic social network modeling and analysis*, (pp. 162–173)
- Mark, N. (1998). Beyond individual differences: Social differentiation from first principles. *American Sociological Review*, (pp. 309–330)
- Marsden, P. V. (1988). Homogeneity in confiding relations. *Social Networks*, 10(1), 57–76
- Mäs, M., Flache, A. & Helbing, D. (2010). Individualization as driving force of clustering phenomena in humans. *PLoS Comput Biol*, 6(10), e1000959

- Mäs, M., Flache, A. & Kitts, J. A. (2014). Cultural integration and differentiation in groups and organizations. In *Perspectives on Culture and Agent-based Simulations*, (pp. 71–90). Springer
- McKerns, M. M., Strand, L., Sullivan, T., Fang, A. & Aivazis, M. A. (2012). Building a framework for predictive science. *arXiv preprint arXiv:1202.1056*
- McKinney, W. (2010). Data structures for statistical computing in python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference*, (pp. 51 – 56)
- McPherson, M., Smith-Lovin, L. & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, (pp. 415–444)
- Miller, A. G., McHoskey, J. W., Bane, C. M. & Dowd, T. G. (1993). The attitude polarization phenomenon: Role of response measure, attitude extremity, and behavioral consequences of reported attitude change. *Journal of Personality and Social Psychology*, 64(4), 561
- O'Donnell, G., Schmitter, P. C., Arnon, C. J. & Lowenthal, A. F. (2013). *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*. JHU Press
- O'Gorman, H. J. (1975). Pluralistic ignorance and white estimates of white support for racial segregation. *Public Opinion Quarterly*, 39(3), 313–330
- Petty, R. E., Wegener, D. T. & Fabrigar, L. R. (1997). Attitudes and attitude change. *Annual Review of Psychology*, 48(1), 609–647
- Pool, G. J., Wood, W. & Leck, K. (1998). The self-esteem motive in social influence: Agreement with valued majorities and disagreement with derogated minorities. *Journal of Personality and Social Psychology*, 75(4), 967
- Prentice, D. A. & Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, 64(2), 243
- Radlauer, D. (2007). Virtual communities as pathways to extremism. *NATO Security through Science Series E Human and Societal Dynamics*, 25, 67
- Ressler, S. (2006). Social network analysis as an approach to combat terrorism: Past, present, and future research. *Homeland Security Affairs*, 2(2)
- Rosenbaum, M. E. (1986). The repulsion hypothesis: On the nondevelopment of relationships. *Journal of Personality and Social Psychology*, 51(6), 1156
- Salzarulo, L. (2006). A continuous opinion dynamics model based on the principle of meta-contrast. *Journal of Artificial Societies and Social Simulation*, 9(1), 13
- Schröder, T. & Thagard, P. (2013). The affective meanings of automatic social behaviors: Three mechanisms that explain priming. *Psychological Review*, 120(1), 255
- Smaldino, P. E. & Epstein, J. M. (2015). Social conformity despite individual preferences for distinctiveness. *Royal Society Open Science*, 2(3), 140437
- Smeaton, G., Byrne, D. & Murnen, S. K. (1989). The repulsion hypothesis revisited: Similarity irrelevance or dissimilarity bias? *Journal of Personality and Social Psychology*, 56(1), 54
- Taber, C. S. & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3), 755–769
- Waskom, M., Botvinnik, O., Hobson, P., Warmenhoven, J., Cole, J. B., Halchenko, Y., Vanderplas, J., Hoyer, S., Villalba, S., Quintero, E., Miles, A., Augspurger, T., Yarkoni, T., Evans, C., Wehner, D., Rocher, L., Megies, T., Coelho, L. P., Ziegler, E., Hoppe, T., Seabold, S., Pascual, S., Cloud, P., Koskinen, M., Hausler, C., kjemmett, Milajevs, D., Qalieh, A., Allan, D. & Meyer, K. (2015). seaborn: v0.6.0 (june 2015). doi:10.5281/zenodo.19108
- Wolf, I., Schröder, T., Neumann, J. & de Haan, G. (2015). Changing minds about electric cars: An empirically grounded agent-based modeling approach. *Technological Forecasting and Social Change*, 94, 269–285
- Wood, W. (2000). Attitude change: Persuasion and social influence. *Annual Review of Psychology*, 51(1), 539–570
- zu Erbach-Schoenberg, E., Bullock, S. & Brailsford, S. (2013). A model of spatially constrained social network dynamics. *Social Science Computer Review*, (p. 0894439313511934)