



Meysam Alizadeh, Alin Coman, Michael Lewis and Claudio Cioffi-Revilla (2014)

## Intergroup Conflict Escalation Leads to More Extremism

*Journal of Artificial Societies and Social Simulation* 17 (4) 4

<<http://jasss.soc.surrey.ac.uk/17/4/4.html>>

Received: 01-May-2012 Accepted: 22-Apr-2014 Published: 31-Oct-2014



### Abstract

Empirical findings in intergroup conflict literature show that individuals that hold beliefs that include differentiation from out-groups become radicalized as intergroup tensions escalate. They also show that differentiation is proportional to tension escalation. In this paper, we present and demonstrate an agent-based model that captures these findings to better understand the effect of perceived intergroup conflict escalation on the average number of emergent extremists and opinion clusters in the population. The proposed model builds on the 2-dimensional Bounded Confidence Model proposed by Huet et al (2008). Results show that the average number of extremists is negatively correlated with intolerance threshold and positively correlated with changes in opinion movement when two agents reject each other's belief. In other words, the number of extremists grows with increasing conflict between groups. We also found that intergroup conflict leads to lower opinion diversity in the population when compared to less conflictual situations.

#### Keywords:

Intergroup Conflict, Opinion Dynamics, Differentiation, Bounded Confidence, Extremism, Radicalization



### Introduction

- 1.1 People are typically more comfortable with those who are similar to themselves or with those who are perceived as in-group members. This phenomenon is called *homophily* and is one of the most widely accepted and empirically proven observations in social psychology (Lazarsfeld & Merton 1954; Burt 1991). Ignorance of or differentiation to people of out-groups can also come from homophilous behavior. This discriminatory and biased behavior of people, based on perceived in-group/out-group membership, can trigger or intensify intergroup conflict.
- 1.2 Several social psychological theories have tried to explain the underlying cognitive mechanisms of differentiation. *Cognitive Dissonance Theory* is an attitude-forming paradigm proposed by L. Festinger (1957). Cognitive dissonance occurs when there is an inconsistency between two or more beliefs that are held simultaneously. Dissonance is "psychologically uncomfortable" and, therefore, causes a person to restore a balanced cognitive state by differentiating from opposing beliefs or other mechanisms. Other theories, such as *Reactance Theory* (Brehm 1966) and *Social Judgment Theory* (Sherif & Hovland 1961), assert that in some cases individuals shift their attitudes in the opposite direction with whom they are interacting.
- 1.3 Differentiation is the main psychological mechanism leading to the emergence of pluralism, polarization, and radicalization in society (Isenberg 1986). The aim of differentiation can be to maintain or to achieve superiority over out-groups (Tajfel & Turner 1986), avoid betraying in-groups (Tajfel & Turner 1986), maintain in-group solidarity (Bulbulia & Sosis 2011), or protect in-group sacred values (Atran & Ginges 2012). However, categorization of persons into different groups by itself does not trigger intergroup conflict. There must be some kind of motivation that activates discriminatory behavior toward out-groups (Fiske 2002). Several social psychological theories have attempted to provide a plausible explanation for the forces that cause intergroup conflicts. These theories vary from interpersonal to social influence-based theories.
- 1.4 Early theories on intergroup conflict explained emergent intergroup conflict as a result of individuals' prejudice and discrimination. Explanations such as the *Theory of Authoritarian Personality* (Adorno et al 1950) and different versions of the *Theory of Frustration, Aggression, and Displacement* (Berkowitz 1962) are well-known instances of this category. The main drawback of these theories is ignorance of social influence over individuals' behavior.

- 1.5 From the social influence point of view, *Realistic Group Conflict Theory* (Sherif 1966) postulates that real conflict of interest between groups triggers intergroup conflict. Another theory concerning cognitive basis in intergroup differentiation is the well-known *Social Identity Theory* (Tajfel 1978), which states that group membership creates in-group feelings that can favor in-group traits, values, and characteristics. However, the theory (Tajfel 1978) and its subsequent formulation, *Self-categorization Theory* (Turner et al 1987), fail to provide a motivational driver for the process of in-group identification and intragroup derogation, particularly that of chronic and long-term identification.
- 1.6 Social pressures to favor one's own group through in-group/out-group bias lead groups to differentiate themselves from one another (Tajfel 1974, 1975; Turner 1975). The amount of differentiation varies by degree, from neglecting out-groups to murderous hostility toward threateningly perceived out-groups. Explicit conflicts of group interests intensify positive attachment to in-groups, leading to antagonistic inter-group relations (Tajfel & Turner 1986). The more intense the inter-group conflict, the more likely that competing groups will act based on their group identification (Tajfel & Turner 1986).
- 1.7 Social psychologists studying intergroup conflict have shown that individuals' out-group attitudes can be predicted by the extent to which they are identified with in-group values. Deschamps and Brown (1983) found that when subjects are primed to categorization, the comparability of groups' roles increases intergroup differentiation. In an attempt to empirically examine the utility of different social psychological approaches for understanding intergroup relations, Brown et al. (1986) conducted an experiment among shop floor workers in a factory. They found that perceived intergroup conflict has a strong positive correlation with differentiation. They also found that the amount of contact is negatively and the strength of group identification is positively correlated with differentiation, but these correlations are weak. To examine the relationship between strength of identification with a political party and the amount of differentiation from other political out-groups, Kelly (1988) showed that "in-group identification and perceived material conflict were consistent predictors of intergroup differentiation along evaluative and affective dimensions but not along the dimension of perceived intragroup homogeneity." Moreover, empirical findings show that when differentiation need is activated in individuals, they generate more negatively evaluated categories compared to when the inclusion need is activated (Leonardelli et al. 2010).
- 1.8 Despite the rich literature on intergroup conflict and intergroup differentiation, there is a need to build computational models to explore the emergent high-level patterns of individuals' opinion formation in presence of intergroup conflict. A plausible direct consequence of the presence of intergroup conflict is the emergence of individuals with radical beliefs. The process is called *radicalization* and has received attention within social psychology from its inception. One important research question in studying extremism is to investigate *why* and *how* individuals with radical beliefs come to believe what they believe. Are they holding radical beliefs because they have interacted with other extremists or could they become radical just by interacting with others who are moderate? What are different plausible pathways or mechanisms for becoming an extremist?
- 1.9 So far, we have discussed how the amount of differentiation can increase in conflict situations (Deschamps & Brown 1983), and that perceived intergroup conflict has strong positive correlation with differentiation (Brown et al. 1986). The aim of this paper is to analyze the effect of perceived intergroup conflict escalation on the average number of emergent extremists and opinion clusters in the population. We develop an agent-based model that links micro-level findings on intergroup conflict and intergroup differentiation literature to macro-level behavior in the population. The proposed model is an extension of the 2-dimensional Bounded Confidence Model with rejection mechanism (2D BCR), as proposed by Huet et al (2008). We randomly assign agents to  $m$  groups and by assuming that the intergroup conflict exists *a priori* in the population, and that agents differentiate more if they perceive more intergroup conflict (Brown et al. 1986), we allow agents to move farther from out-groups' opinion compared to in-group fellows. Before going through the model, we briefly review related agent-based opinion dynamics models in the next section. In section 3 (Intergroup Differentiation Escalation Model), we present our extended model for capturing intergroup conflict escalation and differentiation. We discuss simulation results for various parameters in section 4 (Analysis of Results). Finally, we conclude the paper in section 5.



## Agent-based Opinion Dynamics Models

- 2.1 Agent-based modeling is a fast-growing approach to studying collective behavior in a large number of actors. Many agent-based models have been proposed for opinion dynamics applying social psychology theories. Most existing models are based on the fact that actors coalesce closer following interaction, and homophily is the interpretation offered by these models, which assumes that "attractive forces" exist among individuals holding similar beliefs. Two widely used models in this class are the Bounded Confidence (BC) model (Deffuant et al 2000; Hegselmann & Krause 2002) and the Relative Agreement (RA) model (Deffuant et al 2002).
- 2.2 There are two well-known opinion dynamics under BC models that have been independently proposed by Deffuant, Weisbuch and others (DW), and Hegselmann and Krause (HK) in 2000. They are similar, differing mainly in their communication regimes and slightly in their respective updating mechanisms. While the DW model considers random pairwise encounters at each time step in which agents may compromise or not, the HK model allows agents to communicate with all other agents and adopt the average opinion of those falling within their area of confidence. For a survey of continuous opinion dynamics under bounded confidence models see Lorenz (2007). In this paper, we follow the DW version of the BC model, similar to Huet et al (2008).
- 2.3 In the DW model, agents have continuous belief, ranging from 0 to 1. Agents interact in randomly dyadic encounters and are

allowed to adjust their beliefs if their difference in belief is less than a pre-defined threshold, called "uncertainty." As uncertainty increases, the agents converge to an average belief, but the number of opinion clusters decreases. The DW model fails to create extremists agents because it only considers attractive forces among individuals. While in the DW model, uncertainty thresholds are assumed to be equal for all agents, in the RA model, the authors allow uncertainty to change as a function of time. As a result, after each random paired interaction, both agents' belief and uncertainty are updated and, therefore, influence is no longer symmetric. However, similar to the DW model, the RA model fails to capture the differentiation mechanism in dyadic interaction.

2.4 Several researchers have tried to model the emergence and propagation of extremists in a population using agent-based opinion dynamics models. They attempt to study:

- radicalization by adding some extremist agents in the population with extreme belief but much lower uncertainty threshold (Deffuant et al. 2002)
- asymmetric confidence and biased confidence (Hegselmann & Krause 2002)
- weighting the influence of agents through a Gaussian function with uncertainty as standard deviation (Deffuant et al. 2004)
- examining the effect of network topology (Amblard & Deffuant 2004)
- assigning separate uncertainty thresholds for attraction and rejection (Jager & Amblard 2005)
- incorporating findings from *Self-Categorization Theory* (Salzarulo 2006)
- comparing different continuous opinion dynamics models including the BC model, the Gaussian BC model, the RA model and the Gaussian BC model with interlocutor uncertainty (Deffuant 2006)
- examining the striving for uniqueness among agents (Mäs et al 2010)
- and introducing open- and close-minded agents in the population (Lorenz 2010)

2.5 In an effort to include the rejection mechanism in the DW model, Huet et al (2008) propose a 2-dimensional BC model that allows agents to reject others' beliefs when they are in a dissonant situation. That is, when two agents are close in one attitude and sufficiently far in another, they are in a dissonant state and, therefore, decrease dissonance by shifting away from each other on the attitude on which they are close. By implementing this repulsive force into the BC model, some linear clusters form in attitude borders representing the emergence of extremism in the population. In another work, Huet and Deffuant (2010) introduce a new version of the BC model based on empirical results by Wood et al (1996). While the model is still 2-dimensional, one attitude is considered as a main or important one and the other as a secondary. If two agents are close in primary attitude (difference is less than attraction threshold), they get closer in both attitudes. However, if they are far enough in primary attitude (difference is greater than rejection threshold), then they are in a dissonance state and try to solve it by shifting away in secondary attitude. The results showed that a high rejection threshold leads to the formation of clusters having extreme opinions for the secondary belief.

2.6 The BC model has been extended by Kurmyshev et al (2011), using heterogeneous agents of two types: friendly and partially antagonistic. While the former always reaches others, getting closer to them, the latter exhibits repulsive behavior. In this model radical opinions emerged when there was an equal number of agent types and uncertainty was high.



## Intergroup Differentiation Escalation Model

3.1 In this section we present our agent-based model as an extension of the 2-dimensional Bounded Confidence Model with rejection mechanism proposed by Huet et al (2008). Our modifications are grounded on additional social psychological findings (Deschamps & Brown 1983; Brown et al. 1986; Kelly 1988; Leonardelli et al. 2010). First we describe the Huet et al's (2008) model, followed by our modifications. Consider a set of  $N$  individuals, each with the following characteristics:

1. **Opinion:** a 2-dimensional vector containing  $x_1$  and  $x_2$  represented by real numbers ranging from -1 to +1, reflecting the belief of a node over two different issues. The continuum of opinion can be interpreted as the extent to which agents favor or oppose a given issue.
2. **Uncertainty:** a 2-dimensional vector containing  $u_1$  and  $u_2$  represented by real numbers between 0 and 1 reflecting uncertainties related to  $x_1$  and  $x_2$  respectively.

3.2 At each simulation time step, instead of allowing each agent to interact with all of its neighbors, a pair of individuals is randomly selected to interact and update their beliefs, conditioning the updating on values of beliefs and uncertainties. Suppose agent  $i$  has beliefs  $x_{1i}$  and  $x_{2i}$  with uncertainties  $u_{1i}$  and  $u_{2i}$ , and agent  $j$  has beliefs  $x_{1j}$  and  $x_{2j}$  with uncertainties  $u_{1j}$  and  $u_{2j}$ . For simplicity, all nodes have the same uncertainty  $U$ . Agent  $i$  then compares its beliefs with  $j$ 's and updates. The general rule is that agents approach each other if they are close enough in both beliefs. Otherwise, they may ignore each other or repel and move away. More formally, if

$$|x_{1i} - x_{1j}| \leq U \text{ and } |x_{2i} - x_{2j}| \leq U$$

the two agents' beliefs fall within their bounded confidence interval. Thus, they approach each other, based on the following equations:

$$x_{1i}^{t+1} = x_{1i}^t + \mu (x_{1j}^t - x_{1i}^t) \quad (1)$$

$$x_{2i}^{t+1} = x_{2i}^t + \mu (x_{2j}^t - x_{2i}^t) \quad (2)$$

In equations 1 and 2,  $\mu$  is a constriction factor used to limit convergence velocity, which is assumed to be constant and equal for all agents throughout a simulation run. Another possible state occurs when two agents are close in one belief but far in another:

$$|x_{1i}^t - x_{1j}^t| > U \text{ and } |x_{2i}^t - x_{2j}^t| \leq U$$

Here, two cases arise, depending on whether the difference is less than a certain threshold or not. To represent this, an agent has an "intolerance threshold"  $\delta$ . If the difference is below the predefined threshold, meaning

$$|x_{1i}^t - x_{1j}^t| \leq (1+\delta)U$$

then dissonance is insufficient to trigger rejection. Therefore, the two agents will ignore each other in belief 1 and approach each other in belief 2.

$$x_{1i}^{t+1} = x_{1i}^t \quad (3)$$

$$x_{2i}^{t+1} = x_{2i}^t + \mu (x_{2j}^t - x_{2i}^t) \quad (4)$$

However, if the difference is significant enough, meaning

$$|x_{1i}^t - x_{1j}^t| > (1+\delta)U$$

then conflict will cause agents to feel dissonance, triggering repulsive action, and the two agents will move apart from each other in belief 2 and ignore belief 1. The movement should be large enough to eliminate dissonance

$$x_{2i}^{t+1} = x_{2i}^t - \mu \text{psign}(x_{2j}^t - x_{2i}^t) (U - |x_{2j}^t - x_{2i}^t|) \quad (5)$$

$$x_{1i}^{t+1} = x_{1i}^t \quad (6)$$

where  $\text{psign}(\cdot)$  is similar to the  $\text{sign}$  function, except that it returns +1 if the argument is 0. Moreover, belief values are bound between -1 and +1 based on the following rule:

$$\text{if } |x_{2i}^{t+1}| > 1 \text{ then } x_{2i}^{t+1} = \text{sign}(x_{2i}^{t+1})$$

Since the model is based on *Cognitive Dissonance Theory*, it assumes that when two agents are far on both beliefs, there is no dissonance between them so they simply ignore each other on both beliefs.

- 3.3 Given our research questions, we modified the 2D BCR model in three ways. First, we made the model more realistic by having agents interact in a social network structure with their immediate neighbors. (In the 2D BCR model agents are allowed to interact with any random agent in the population.) Second, to capture the group identification mechanism, we randomly assign all agents to  $m$  different groups. Third, we modify the belief updating rule to incorporate empirical findings from the social psychology literature. In general, findings show that escalation of perceived intergroup tensions leads individuals to greater differentiation from out-groups (Deschamps & Brown 1983; Brown et al. 1986). In our model, we assume that some background tension between groups always exists *a priori* in the population. Moreover, since the findings hold true for "differentiating beliefs," we assume that  $x_1$  and  $x_2$  are differentiating beliefs. An example of a differentiating belief is a "moral belief," including religious beliefs. The differentiation escalation mechanism occurs when two encountering agents belong to different groups. We model the increase in opinion movement by including an *intergroup differentiation escalation* coefficient  $\beta$  ( $\beta > 1$ ) in equation 5:

$$x_{2i}^{t+1} = x_{2i}^t - \beta \mu \text{psign}(x_{2j}^t - x_{2i}^t) (U - |x_{2j}^t - x_{2i}^t|) \quad (7)$$

where  $\beta$  is  $> 1$ . The second term on the right of equation 5 determines the distance that agents move apart from each other in a dissonant situation. By multiplying the coefficient  $\beta$  ( $\beta > 1$ ) in this quantity, we increase the movement and let the agents begin to drift farther apart in their beliefs, compared to the 2D BCR model. As tension escalates between groups, we can increase the value of  $\beta$ , making agents move farther away. When there is no tension between groups,  $\beta$  is set to 1. Based on these modifications, we expect to see more extremists in the population, as we increase the intergroup differentiation escalation coefficient  $\beta$ . In other words, the number of emergent extremists is a function of  $\beta$  and  $\delta$

$$N^t(e) = f^t(\beta, \delta) \quad (8)$$

where the  $N^t(e)$  represents the number of extremists at time step  $t$ ,  $f^t$  is the interaction situation at time step  $t$  (attraction or

rejection);  $\beta$  is the intergroup differentiation escalation coefficient ( $\beta > 1$ ), and  $\delta$  is the intolerance threshold. Table 1 compares the key features of the proposed model with other existing agent-based opinion dynamics models in the literature.

**Table 1:** Light Docking of Agent-based Opinion Dynamics Models

Variables	BC (Deffuant et al 2000)	RA (Deffuant et al 2002)	Salzarulo (2006)	Huet et al (2008)	Differentiation Escalation
Supporting Theory	Cognitive Dissonance	Cognitive Dissonance	Self-Categorization	Cognitive Dissonance	Brown et al (1986)
Dimension	1	1	n	2	2
Uncertainty	Constant	Dynamic	N.A.	Constant	Constant
Communication Regime	Dyadic	Dyadic	Group	Dyadic	Dyadic
Rejection Mechanism	×	×	✓	✓	✓
Differentiation Escalation Mechanism	×	×	×	×	✓
Initial Groupings	×	×	✓	×	✓



## Simulation Results

- 4.1 In this section, we present and analyze our agent-based simulation results. First we show the general behavior of the model and compare it to the original Huet et al (2008) model. Second, we show virtual experimental results on the effect of key factors on the number of emergent extremists in the population.

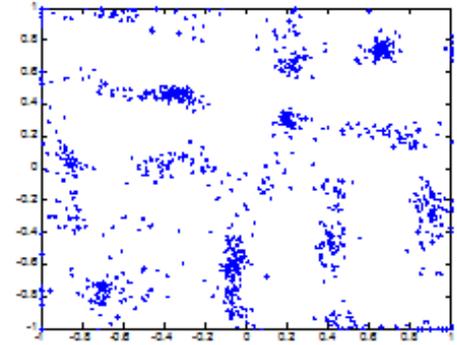
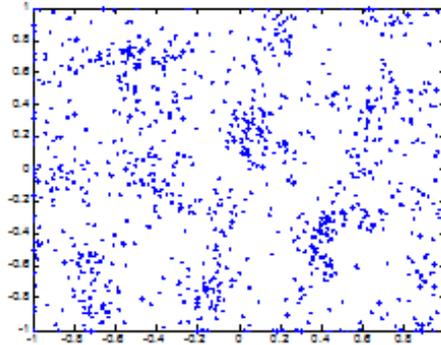
### General Comparison with the 2D BCR Model

- 4.2 We consider a population of 1,000 agents, each having two beliefs. Agents' initial beliefs are randomly assigned using a uniform distribution between -1 and 1. Uncertainty  $u_1$  and  $u_2$  are assumed to have equal values held constant throughout the simulation. The assumption is that  $x_1$  and  $x_2$  are differentiating beliefs and agents are in a conflict escalation situation (i.e.  $\beta > 1$ ). Since our main focus is to test the effect of intergroup differentiation escalation, we omit the network structure in this section and set the same communication regime as in the 2D BCR model.
- 4.3 Figure 1 compares the evolution of opinions between the Huet et al (2008) model and our modified model. In each of the figures, the two axes represent the beliefs that are bounded between -1 and 1 and each dot represents an agent's opinion. Our model's assumption is that both beliefs are differentiating and there is tension between groups. All other parameters are the same. Therefore, we can interpret results as changes in a population's belief as such beliefs become differentiating and tensions arise between groups.
- 4.4 After some time, several stable equilibria emerge and opinion clusters form around them, similar to the 2D BCR model. Two interactive social forces cause the creation of the clusters. One force affects agents that are close in both beliefs, making them coalesce and form groups of like-minded agents. The other force affects those who hold similar belief in one dimension and dissimilar enough in another, separating and pushing them away from each other. Consequently, after some time, some dominant equilibria emerge in the population (Huet et al 2008). These "meta-clusters" are permanent and do not change their position on the figure.

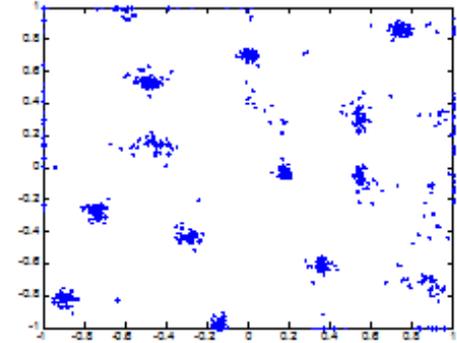
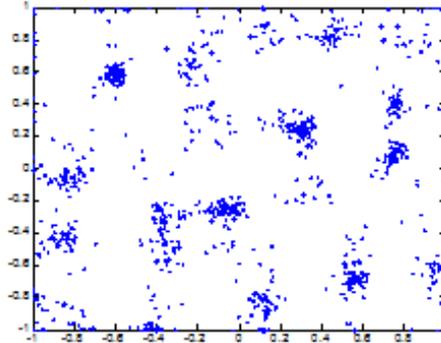
## Differentiation Escalation Model

## 2D BCR Model

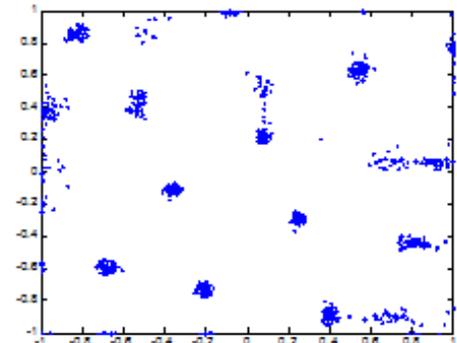
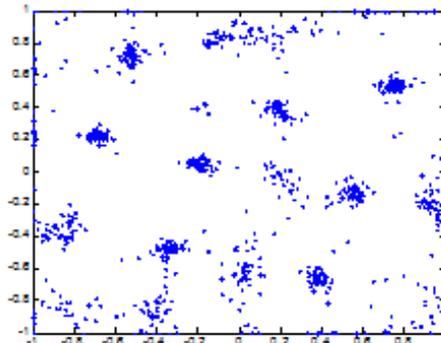
T=100,000



T = 200,000



T = 300,000



**Figure 1.** Comparison of our differentiation escalation model with the 2D BCR model ( $U=0.2$ ,  $m=3$ ,  $\mu=0.3$ ,  $\delta=1$ ,  $\beta=1.5$ ). Both axes represent individuals' opinion.

- 4.5 An interesting difference between these results is that the number of emergent opinion clusters is less in our differentiation escalation model than in the earlier 2D BCR model. Our model uses Deffuant's (2006) and Huet et al's (2008) algorithm to compute the number of clusters. That is, we define a minimum distance  $\varepsilon$  between agents' beliefs when assigning agents to the same cluster. In practice, we used  $\varepsilon = 0.15$  and neglected clusters of size smaller than %1 of the population size. We ran the model 25 times and recorded the average number of clusters.
- 4.6 Results show that while the 2D BCR model produces 19 opinion clusters with at least 10 members, the differentiation escalation model produces 16 opinion clusters. The difference can be explained by the role of  $\beta$  which increases fluctuations of individuals within clusters, thereby increasing the instability of the clusters and the likelihood of clusters coalescing or, in rare cases, clusters exploding. This mechanism eventually produces fewer opinion clusters, implying that, in the presence of escalating tensions and differentiating opinions, not only do equilibrium points change, but we also see less opinion diversity in the society. Later, we explore this phenomenon in greater detail (subsection Analyzing the Effect of  $\beta$ ).

### Virtual Experiments

- 4.7 In this section, we conduct a virtual experiment for testing the effect of key variables on the average number of emergent extremists in the population. We define an agent as extremist when the absolute value of at least one of its beliefs is equal or greater than 0.9. Previous researchers have examined the effect of intolerance thresholds (Deffuant 2006; Huet et al. 2008), the initial uncertainty of moderates (Deffuant 2006), the initial proportion of extremists (Deffuant 2006), level of uncertainty (Huet et al 2008), and network topology (Amblard & Deffuant 2004; Weisbuch 2004) on emergent dynamics and number of extremists. To avoid the analysis being too complicated and lengthy, we only vary the newly introduced parameters and the intolerance threshold variable, controlling the convergence parameter, level of uncertainty, and number of agents. We tested for values of

intergroup differentiation escalation  $\beta$ , intolerance threshold  $\delta$ , and number of initial groupings  $m$ . We also examined the effect of a random social network by adding an indicator variable that is 1 when agents interact with immediate neighbors using an Erdős-Rényi random network with connectivity probability of  $p = 0.05$ , and 0 when there is no network structure. Each combination was run 25 times. Results are reported in Table 2.

- 4.8 The ANOVA table contains the sources of variation, degrees of freedom, sum of squares (SS), mean square (MS) values,  $F$ -ratio test statistics, and the corresponding significance levels ( $p$ -values). In general, the higher the  $F$ -ratio value, or the smaller the probability, the more important the corresponding factor. Table 3 presents results of the ANOVA test for the average number of extremists. Results show significant difference between levels of intergroup differentiation escalation  $\beta$  ( $F(2, 1296) = 3.002, p = 0.050$ ) and intolerance threshold  $\delta$  ( $F(2, 1296) = 1229.465, p = 0.000$ ). When examining the initial number of groupings  $m$ , we observed a significant difference ( $F(2, 1296) = 24.525, p = 0.000$ ). Finally, the absence or presence of a random network is also significant ( $F(1, 1296) = 13.694, p = 0.000$ ).

**Table 2:** Experimental Design

Independent Variables	No. of Test Cases	Values Used
Beta	3	1, 1.5, 2
Intolerance threshold	3	1, 1.5, 2
Number of initial groupings	3	3, 5, 7
Social Network	2	Random ( $p = 0.05$ ), No network
Control Variables	No. of Test Cases	Values Used
Uncertainty ( $U$ )	1	0.2
Convergence Parameter ( $\mu$ )	1	0.3
Initial Belief Distribution	1	Random
Maximum Iteration	1	500,000
Number of Agents	1	1000
Number of Runs	1	25
Dependent Variables	No. of Test Cases	Values
Average Number of Extremists	-	-

- 4.9 The ANOVA table also reveals that there are significant interactions between intergroup differentiation escalation  $\beta \times$  intolerance threshold  $\delta$  ( $F(4, 1296) = 22.577, p = 0.000$ ), the number of initial groupings  $m \times \beta$  ( $F(4, 1296) = 2.558, p \leq 0.05$ ),  $\beta \times$  random network existence ( $F(2, 1296) = 8.274, p = 0.000$ ),  $\delta \times m$  ( $F(4, 1296) = 5.892, p = 0.000$ ), and  $m \times$  random network existence ( $F(2, 1296) = 9.239, p = 0.000$ ). This means that the simultaneous influence of independent variables on the average number of emergent extremists is not additive. That is, the relationship between each of the interacting variables and the dependent variable depends on the value of the other interacting variable. Further analysis of the effect of independent variables is presented in the following three sub-sections.

**Table 3:** ANOVA Table for Tests of Between-Subjects Effects

Source	SS	Degrees of Freedom	MS	F	P-Value
Beta	2505.975	2	1252.987	3.002	.050
Delta	1026453.255	2	513226.627	1229.465	.000
m	20475.539	2	10237.770	24.525	.000
Network	5716.507	1	5716.507	13.694	.000
Beta * Delta	37698.167	4	9424.542	22.577	.000
Beta * m	4271.563	4	1067.891	2.558	.037
Beta * Network	6907.471	2	3453.736	8.274	.000
Delta * m	9838.776	4	2459.694	5.892	.000
Delta * Network	880.164	2	440.082	1.054	.349
m * Network	7713.524	2	3856.762	9.239	.000
Beta * Delta * m	4398.121	8	549.765	1.317	.231
Beta * Delta * Network	32612.404	4	8153.101	19.531	.000
Beta * m * Network	524.271	4	131.068	.314	.869

Delta * m * Network	4917.298	4	1229.324	2.945	.019
Beta * Delta * m * Network	8721.520	8	1090.190	2.612	.008
Error	541000.720	1296	417.439		
Total	67483898.00	1350			

Dependent Variable: Number of Extremists  
R Squared = .684 (Adjusted R Squared = .672)

#### Analyzing the Effect of $\beta$

4.10 Now we analyze the effect of intergroup differentiation escalation  $\beta$  from two perspectives: 1) the effect of  $\beta$  on the average number of emergent extremists in the society, and 2) the effect of  $\beta$  on the number of opinion clusters. Figure 2 compares the number of extremists in the original 2D BCR model and our intergroup differentiation escalation model resulting from 25 runs. We show the number of extremists when interactions are affected by a random network and when they are not. In general, results show that our model produces more extremists compared to the 2D BCR model, independent of the presence of a network structure. In addition, the average number of extremists increases monotonically up to  $\beta = 1.8$ , decreasing afterward as in a convex function.

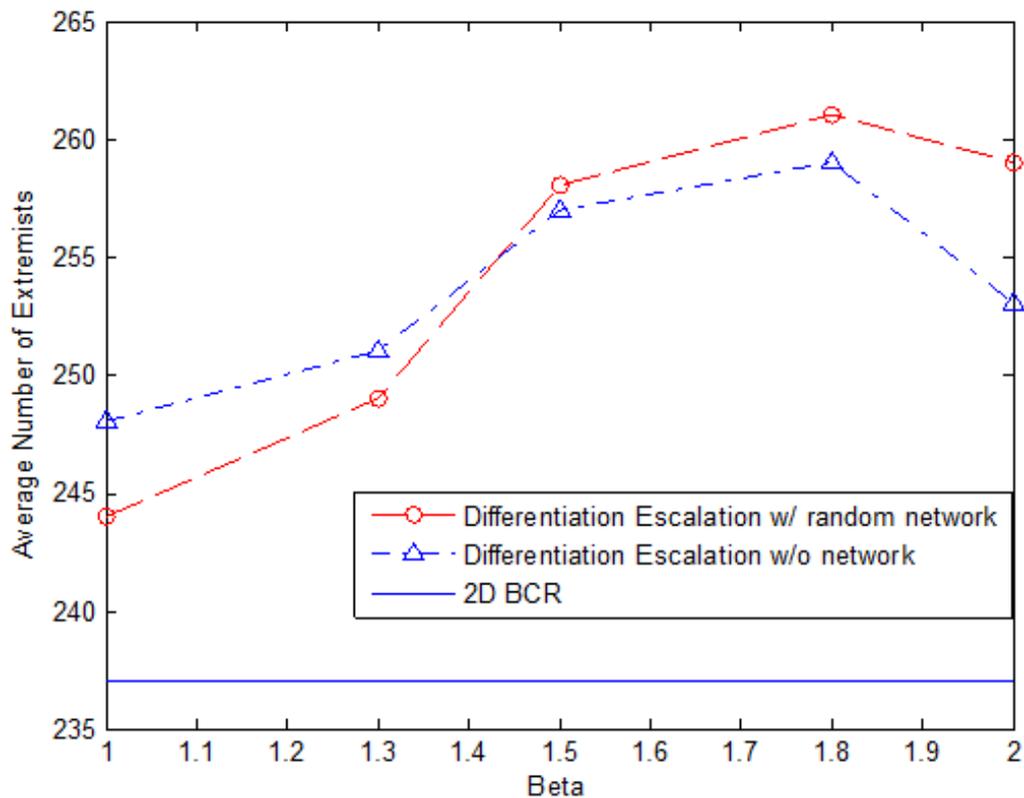
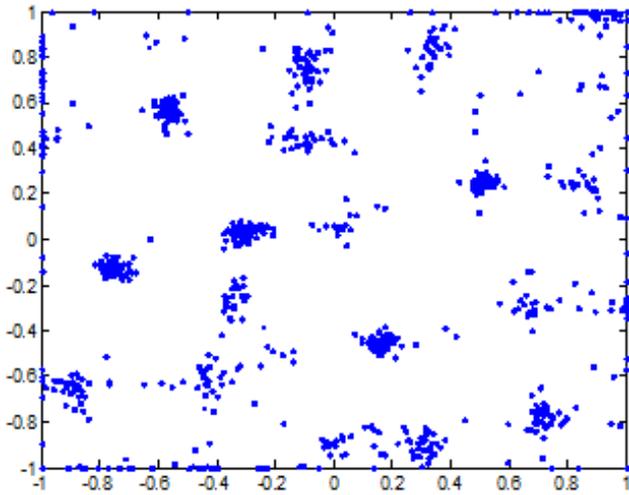


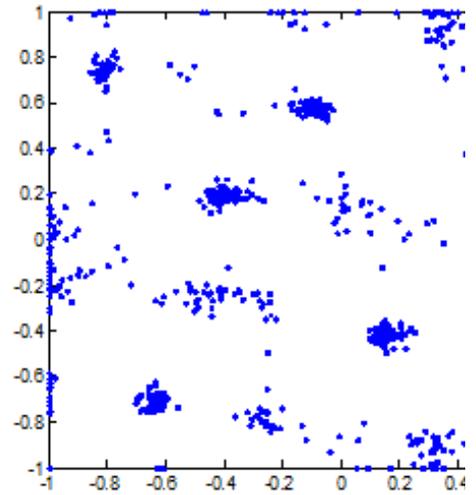
Figure 2. Comparing the effect of  $\beta$  on the average number of extremists between the 2D BCR and differentiation escalation models ( $U = 0.2$ ,  $m = 3$ ,  $\mu = 0.3$ ,  $\delta = 1$ , iterations=500,000).

4.11 The reason for this convex behavior could be traced to the amount of movement that agents undertake when encountering each other. By repelling agents when they are in a dissonant condition, more agents tend to reach the borders of opinion and become extremists. This is an interesting result because it shows that just by changing the social interaction regime, more extremists emerge in the population. This supports the claim of *Self-Categorization Theory* (Turner et al. 1987) that extremists are the product of simple social interactions and that extremism is more likely to be an emerging phenomenon rather than an intrinsic feature of individuals. That is, even without having individuals interacting with extremists in the society, they still can become extremists themselves just by interacting with those who hold less extreme beliefs (Urbig & Malitz 2007). These results also are in accordance with Isenberg's (1986) argument that group polarization is a function of "persuasive arguments" and "social comparison" and, therefore, the existence of extremists is not necessary for polarization and radicalization to occur.

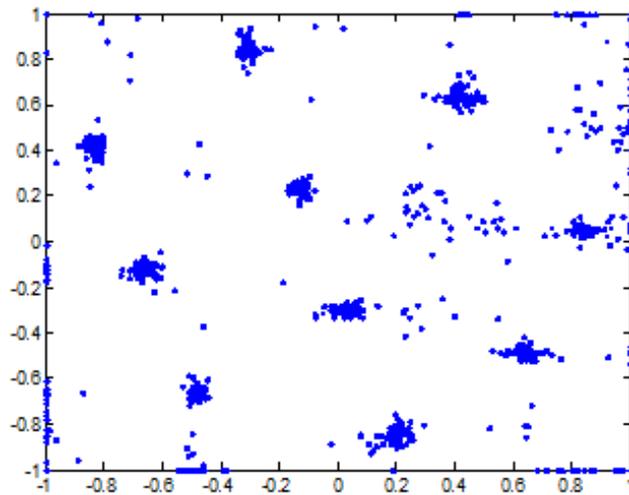
$\beta = 1.2$



$\beta = 1.5$



$\beta = 2$



$\beta = 2.5$

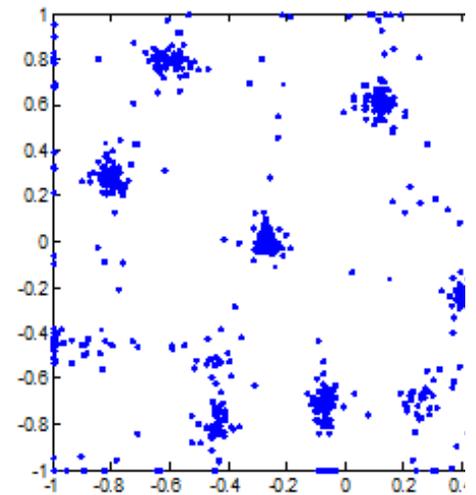
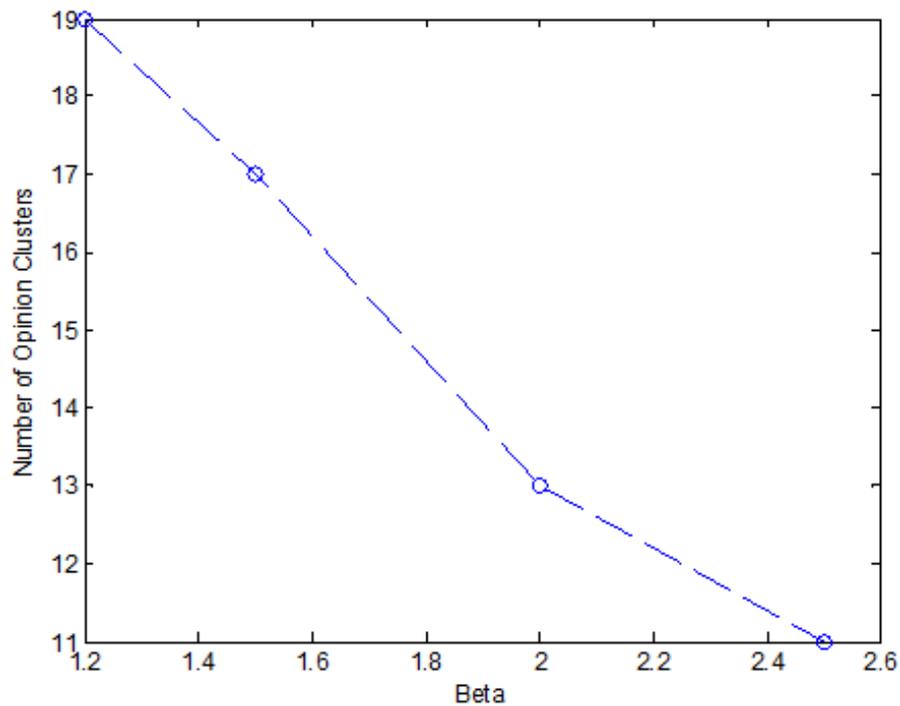


Figure 3. Effect of  $\beta$  on the number of extremists and clusters ( $U=0.2$ ,  $\mu=0.3$ ,  $\delta=1$ , random network, iterations=500)



**Figure 4.** Analyzing the effect of  $\beta$  on the average number of emergent opinion clusters ( $U=0.2$ ,  $\mu=0.3$ ,  $\delta=1$ , random network, iterations=500,000)

- 4.12 In an earlier section (General Comparison with the 2D BCR Model), we discussed how increasing the amount of movement at the time of differentiation increases the instability of opinion clusters by increasing the likelihood of clusters merging or disintegrating, which, in turn, decreases the final number of opinion clusters. Here we look more closely and test the effect of increasing  $\beta$  on the number of opinion clusters. Figure 3 shows simulation results for different values of  $\beta$  and Figure 4 shows the corresponding average number of opinion clusters obtained by Huet et al's (2008) algorithm. We run the algorithm for a maximum number of 500,000 iterations and replicate 25 times. Both figures suggest that the number of emergent opinion clusters decreases with  $\beta$ . Indeed, as we push agents to shift their beliefs farther away from each other, the equilibria relocate and distance between new clusters increases. This suggests that intergroup differentiation escalation reduces opinion diversity in the society and contributes to producing a radicalized population.

#### *Analyzing the Effect of Intolerance Threshold*

- 4.13 The intolerance threshold plays an important role in agents' selection process of whether to become attracted to or differentiated from others. Figure 5 shows simulation results for four different values of intolerance threshold  $\delta$  ( $\delta=1, 1.5, 2$ , and  $2.5$ ), while holding all other variables constant. Observationally, we can see that the number of opinion clusters strikingly increase as the rejection condition is restricted. Figure 6 illustrates the corresponding average number of opinion clusters (iterations = 500,000; replications = 25), where an increase in  $\delta$  from 1.0 to 2.5 increases the average number of opinion clusters from 16 to 29. Thus, we conclude that in intergroup differentiation escalation situations, restricting the differentiation mechanism would increase opinion diversity in population. The same trend was found in the 2D BCR model by Huet et al (2008).

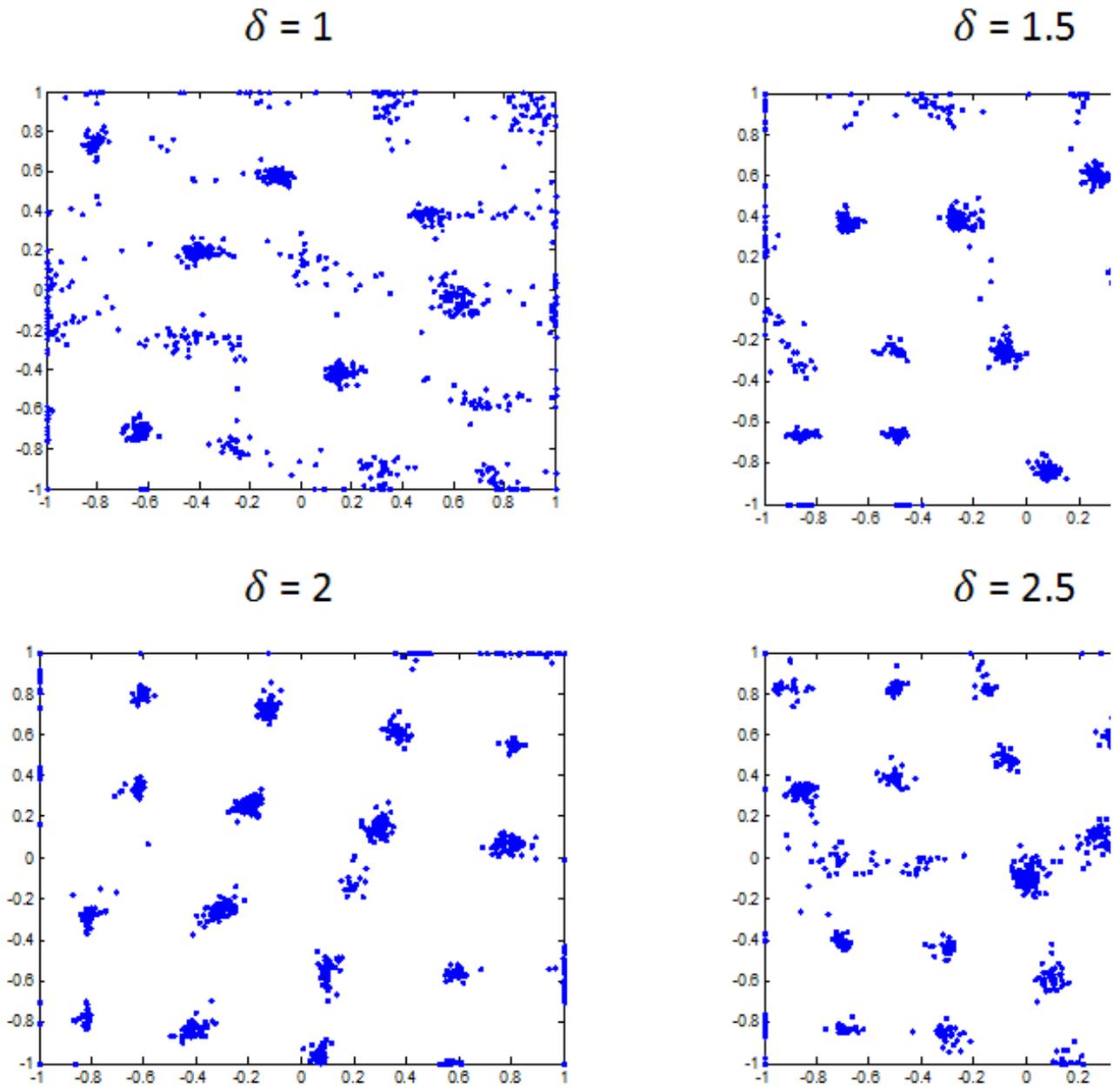
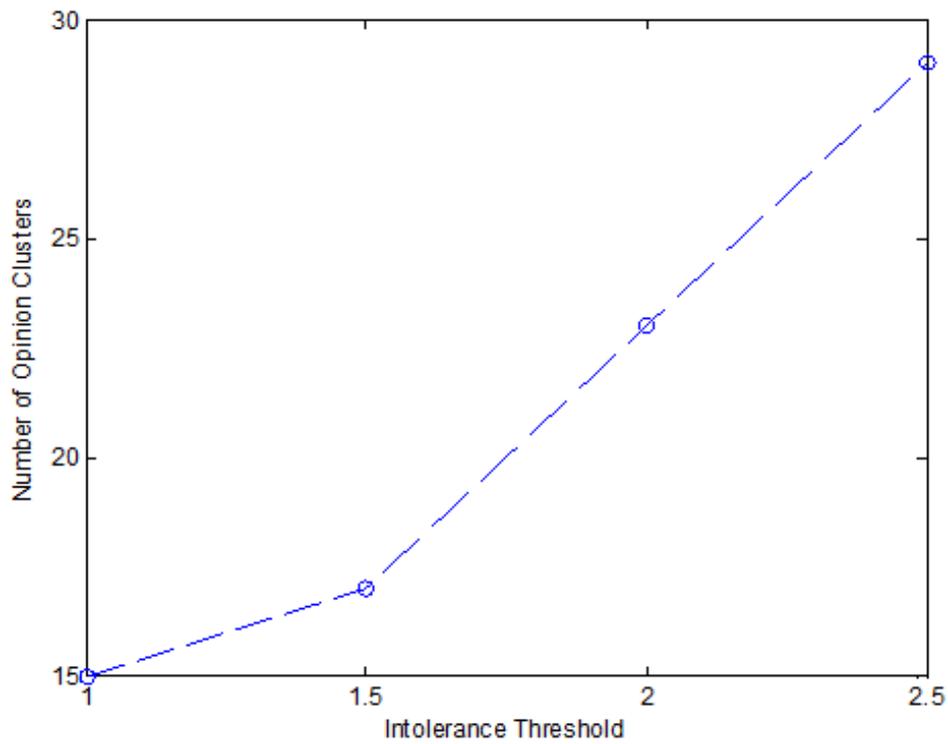
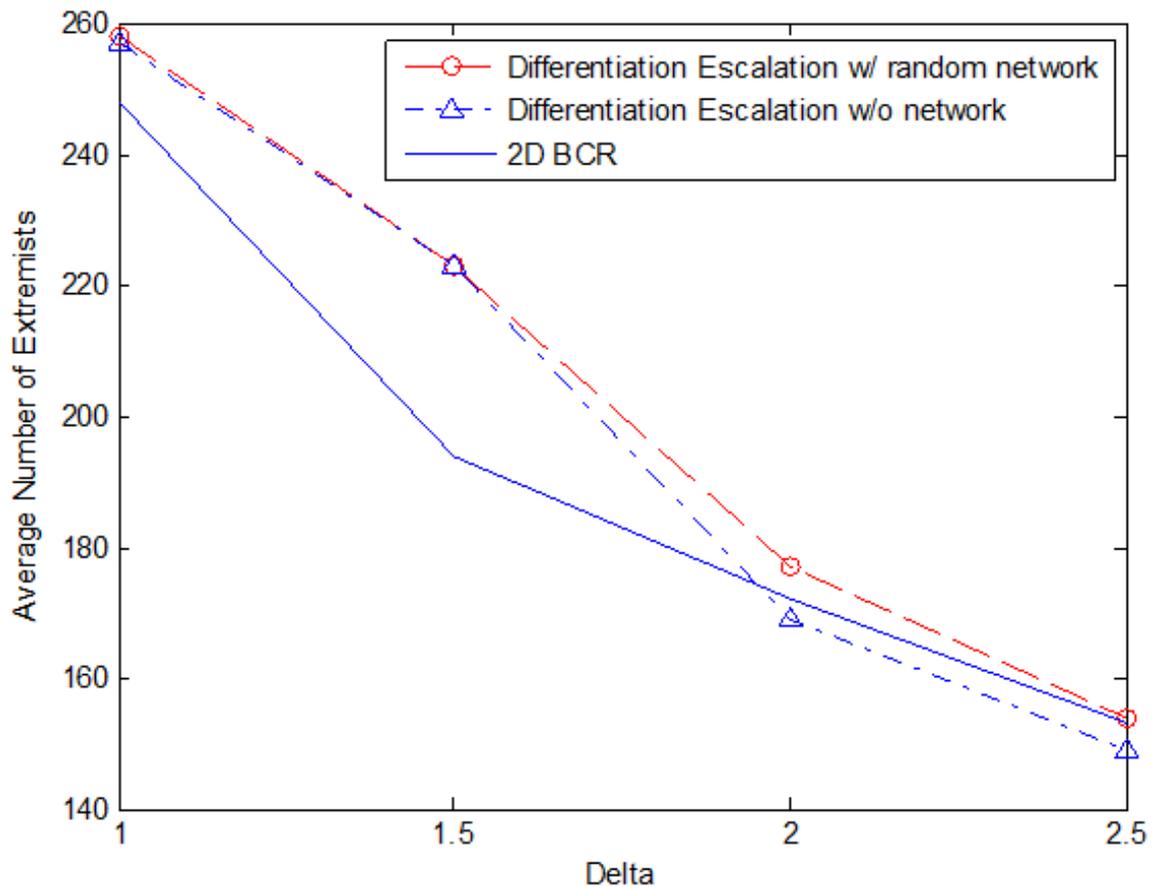


Figure 5. Effect of intolerance threshold  $\delta$  on opinion dynamics ( $U=0.2$ ,  $\mu=0.3$ ,  $\beta=1.5$ , random network, iterations= $\ell$ )



**Figure 6.** Analyzing the effect of intolerance threshold  $\delta$  on average number of emergent opinion clusters ( $U=0.2$ ,  $\mu=0.3$ ,  $\beta=1.5$ , random network, iterations=500,000)

**4.14** ANOVA test results in Table 3 demonstrated that intolerance threshold  $\delta$  has a significant effect on generating extremists. Figure 7 illustrates the average number of extremists when the intolerance threshold  $\delta$  increases from 1 to 2.5. We observe that the number of emerging extremists monotonically decreases by increasing the intolerance threshold. This is due to the fact that by restricting the rejection conditions, people tend to cluster closer to each other or at most ignore each other's opinion rather than rejecting and shifting away. Results also show that, except for  $\delta = 2.5$ , at other corresponding intolerance threshold values, the differentiation escalation model (with or without a random network) produces more extremists.



**Figure 7.** Effect of intolerance threshold  $\delta$  on the average number of extremists ( $U=0.2$ ,  $\mu=0.3$ ,  $\beta=1.5$ ,  $m = 3$ , iterations=500,000)

- 4.15 We now examine the effect of the number of initial groupings  $m$  on the average number of emergent extremists and number of final opinion clusters. Figure 8 shows the final distribution of the opinions after 500,000 iterations for different numbers of initial groupings. Figure 9 shows the corresponding number of final opinion clusters. Both figures suggest that the average number of final opinion clusters slightly decreases with increasing values of  $m$ . This could be explained by the fact that increasing the number of initial groupings increases the differentiation among agents, which leads to a decrease in the number of opinion clusters.
- 4.16 Table 3 showed that changes in the initial number of groupings is highly correlated with the number of emergent extremists in the population. Figure 10 shows the corresponding average number of extremists for various values of initial number of groupings  $m$ , with or without both random network interactions. It can be seen that the average number of extremists increases with the number of initial groupings  $m$  and this effect is independent of neighboring interactions. The reason is that, on the one hand, opinions are bounded, and on the other hand, differentiation increases with a larger number of pre-defined groups in the initial population, so more agents are pushed to the extreme limits of opinions. Results also reveal that random network interactions generate more extremists in this parameter configuration.

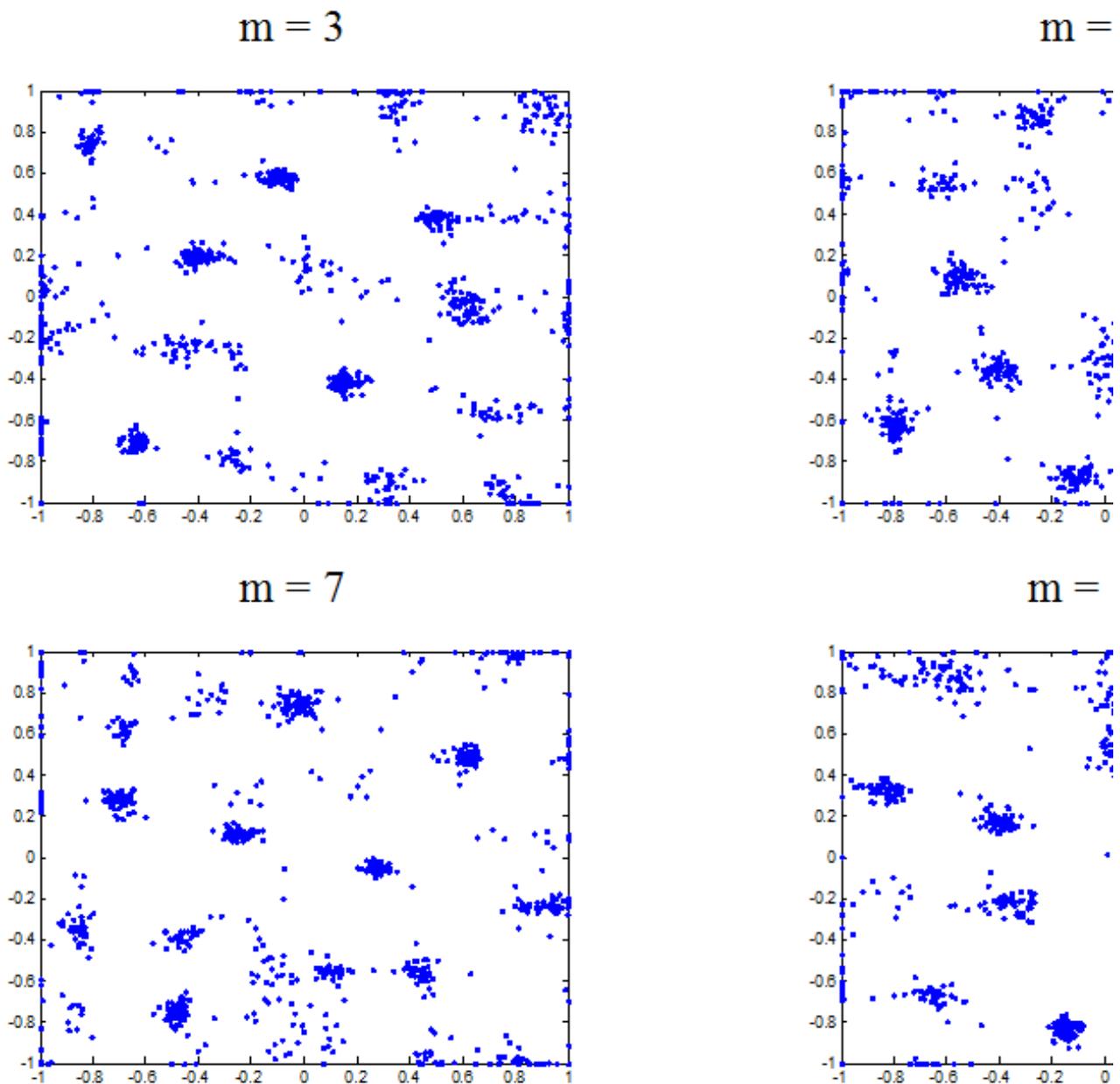


Figure 8. Analyzing the effect of number of initial groupings on opinion dynamics ( $\mathcal{U}=0.2$ ,  $\mu=0.3$ ,  $\beta=1.5$ , random network)

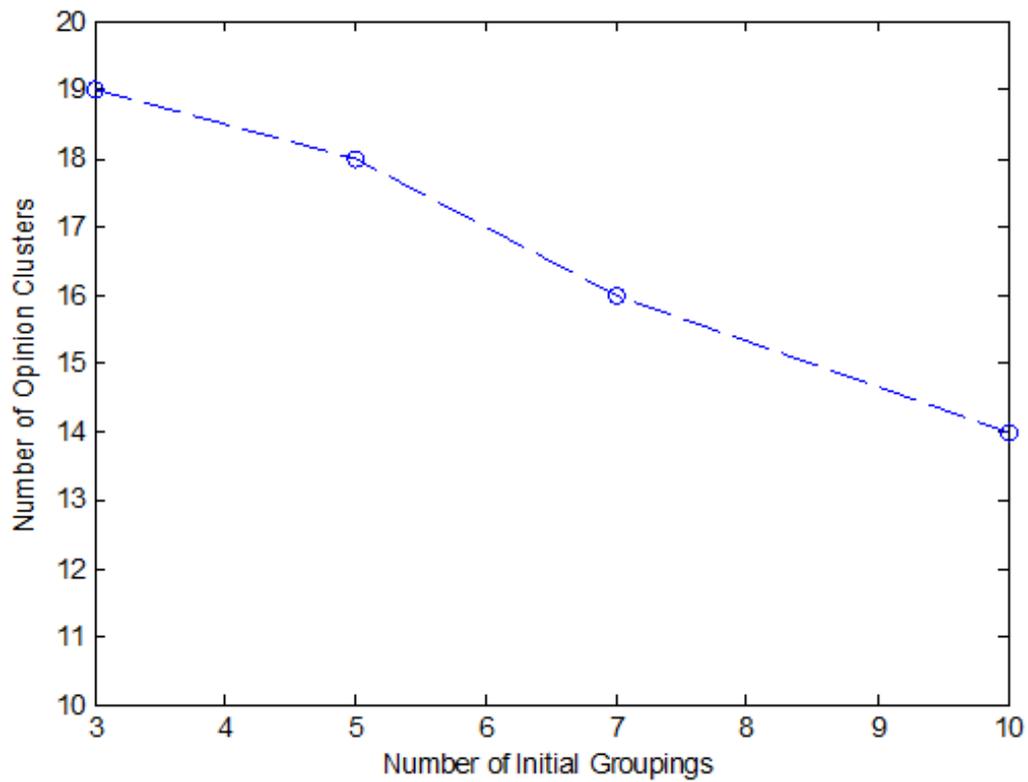


Figure 9. Analyzing the effect of initial number of groupings on average number of emergent opinion clusters ( $U=0.2$ ,  $\mu=0.3$ ,  $\delta=1.5$ ,  $\beta=1.5$ , random network, iterations=500,000)

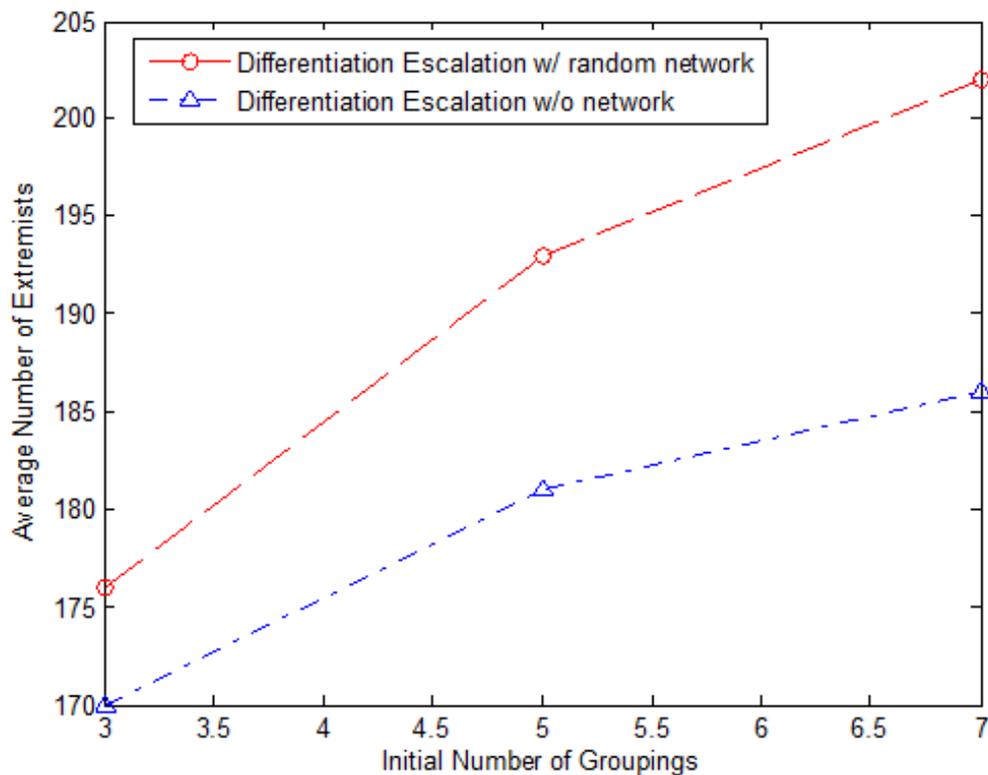


Figure 10. Effect of initial number of groupings  $m$  on average number of extremists ( $U=0.2$ ,  $\mu=0.3$ ,  $\beta=2$ ,  $\delta=1.5$ , iterations=500,000)



## Discussion and Conclusion

- 5.1 Assuming that there is an interest in controlling and mitigating the spread of radical beliefs in society, it is necessary that evidence-based policies (Lum & Kennedy 2012; Cioffi-Revilla 2012) be based on deep understanding of different conditions under which individuals develop radical beliefs. For example, a valid inference from *Self-categorization Theory* is that the existence of extremists in society is an emergent phenomenon, not an intrinsic attribute of individuals. That is, extremists are not the cause of extremism; rather, they are the product of social interactions, which is also consistent with current process-oriented

theories of criminology (Cioffi 2012; McCauley & Moskaleiko 2005) rather than earlier attribute-based theories.

- 5.2 A body of literature has explored the idea that individuals' beliefs that generate differentiation from out-groups create radicalization as intergroup tensions escalate (Deschamps & Brown 1983; Brown et al. 1986; Kelly 1988; Leonardelli et al. 2010). Their results show that conflict triggers identification and differentiation dynamics. Moreover, they find that the more intense the intergroup tensions, the more identification and, hence, the more differentiation occurs between individual's opinions. That is, participants involved in intense conflict identified more with their in-group than participants in lower conflict conditions. The latter also identified more with their in-group than participants in collaborative relations. The main idea is that individuals may not follow rational decision making when they engage in intense conflict because they may consider some issues as sacred values and become unwilling to compromise (Atran & Ginges 2012).
- 5.3 In this study, we showed how and why the collective behavior of individuals produces tensions that trigger in-group opinions between groups of people. Specifically, we tested the effect of conflict escalation on the average number of extremists that emerge as a result of social interactions. We developed our model based on two critical assumptions. First, we assumed that some tensions exist *a priori* in any given society. Second, we assumed agents generate differentiation via normal social interactions. Our agent-based belief differentiation model is a modified version of the 2-dimensional Bounded Confidence with rejection mechanism model proposed by Huet et al (2008). We randomly assigned agents to  $m$  groups and allowed them to move away from each other when they belonged to different groups. We captured this increase in opinion rejection by introducing the intergroup differentiation escalation coefficient  $\beta$ . We also included a social network structure that constrains agents to only interact with their neighbors. We ran the model for various ranges of parameters and observed significant differences in number of opinion clusters, change in equilibria, and levels of consensus. We also conducted virtual experiments to systematically test the effect of intergroup conflict escalation on the average number of emerging extremists.
- 5.4 ANOVA test results support the hypothesis that the average number of extremists is a function of individual-based intolerance threshold and amount of attitude shifts when two agents reject each other's beliefs. In other words, the increase in the amount of opinion differentiation eventually leads to more extremism. By varying the intergroup differentiation escalation coefficient between 1 and 2, we observe more agents at the fringes of opinions, meaning the emergence of more extremists in the population. The reason can be traced to the role of  $\beta$ , which increases the amount of movement when agents reject each other. In our model, in rejection situations, agents are pushed farther away from their cluster by other clusters, as compared to the earlier 2D BCR model. Since the beliefs are bounded between -1 and 1, this increase in the amount of movement increases the distance between clusters which in turn reduces the number of emerging clusters. Results also show that the modified model produces fewer opinion clusters, implying that, in the presence of tension and differentiating beliefs, not only the equilibria change, but also we observe less opinion diversity in the society.
- 5.5 Another interesting result coming from our virtual experiments concerns the role of social network structure on opinion dynamics. Through ANOVA tests, we found that the presence of a random network does *not* have a statistically significant effect on the average number of emerging extremists. Results also show that, as equilibria change as a result of restricting individuals' interactions to a social network, so opinion clusters occasionally relocate.
- 5.6 Analyzing the effect of intolerance threshold  $\delta$  on the dynamics of opinion reveals that it has an inverse correlation with the average number of emerging extremists. That is, as we restricted the rejection condition by expanding the intolerance threshold, on average we see less extremists at the end of the simulation. This is mainly due to the fact that people in a tolerant society are more likely to be attracted to (or at worst ignore) each other's opinion, which mitigates differentiation. We also find that a larger number of opinion clusters emerge as  $\delta$  increases, meaning more opinion diversity in the society.
- 5.7 Future research should test the effect of control variables. In particular, analyzing the effect of different initial belief distributions would be of interest. In this study, we assume that initial beliefs are uniformly assigned to agents. Other distributions can be used, such as normal, or Weibull, to measure their impact on emergent dynamics (Cioffi-Revilla 2014: 152-161; Alizadeh & Cioffi-Revilla 2014). Another possible direction for extending this research is to create a social network structure based on the homophily effect, and then assign initial beliefs accordingly. That is, people who are connected to each other in a social network tend to have characteristics and beliefs that are more similar than those who are not part of the network. Finally, exploring the effect of conflict escalation and the presence of differentiating beliefs on the co-evolution of opinion dynamics, and the structure of the social network structure at the group level, would shed light on dynamics of group membership and the emergence of radical groups.



## Note

The Python code of the model is available in CoMSES Net Computational Model Library at <http://www.openabm.org/model/4154/version/2/view>.



## References

- ADORNO, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The Authoritarian Personality*. Oxford, England: Harpers.
- ALIZADEH, M., & Cioffi-Revilla, C. (2014). Distributions of Opinion and Extremist Radicalization: Insights from Agent-Based Modeling. In *Proceedings of the 6th International Conference on Social Informatics (SoCInfo14), Barcelona, Spain, 10-13 November*, Heidelberg: Springer.
- 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.
- ATRAN, S., & Ginges, J. (2012). Religious and Sacred Imperatives in Human Conflict. *Science*, 336(6083), 855–857.
- BERKOWITZ, L. (1962). *Aggression: A social psychological analysis*. New York, NY, US: McGraw-Hill.
- BREHM, J.W. (1966). *A theory of psychological reactance*. New York Academic Press.
- BROWN, R., Condor, S., Mathews, A., Wade, G., & Williams, J. (1986). Explaining intergroup differentiation in an industrial organization. *Journal of Occupational Psychology*, 59(4), 273–286.
- BULBULIA, J. and Sosis, R. (2011). Signalling theory and the evolution of religious cooperation. *Religion*, 41(3), 363-388.
- BURT, R. S. (1991). Age as a Structural Concept. *Social Networks*, 19, 355–373.
- CIOFFI-REVILLA, C. (2010). A Theory of Individual Radicalization: Modeling Complexity in Terrorism and Political Violence. Working Paper, Center for Social Complexity, George Mason University, Fairfax, Virginia, USA.
- CIOFFI-REVILLA, C. (2012). A Complexity Method for Assessing Counterterrorism Policies. In Lum., C & Kennedy, L.W. (Eds.), *Evidence-Based Counterterrorism Policy*. New York and Heidelberg: Springer. Pp. 151-165.
- CIOFFI-REVILLA, C. (2014). *Introduction to Computational Social Science: Principles and Applications*. London and Heidelberg: Springer.
- DEFFUANT, G., Neau, D., Amblard, F., & Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3, 87–98.
- 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. <http://jasss.soc.surrey.ac.uk/5/4/1.html>
- DEFFUANT, G., Amblard, F., & Weisbuch, G. (2004). Modelling group opinion shift to extreme: the smooth bounded confidence model. *Arxiv preprint cond-mat/0410199*
- DEFFUANT, G. (2006). Comparing Extremism Propagation Patterns in Continuous Opinion Models. *Journal of Artificial Societies and Social Simulation*, 9(3) 8. <http://jasss.soc.surrey.ac.uk/9/3/8.html>
- DESCHAMPS, J. C., & Brown, R. (1983). Superordinate goals and intergroup conflict. *British Journal of Social Psychology*, 22(3), 189–195.
- FESTINGER, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- FISKE, S. T. (2002). What We Know Now About Bias and Intergroup Conflict, the Problem of the Century. *Current Directions in Psychological Science*, 11(4), 123–128.
- HEGSELMANN, R. & Krause, U. (2002). Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2. <http://jasss.soc.surrey.ac.uk/5/3/2.html>
- HUET, S., Deffuant, G., & Jager, W. (2008). Rejection mechanism in 2D bounded confidence provides more conformity. *Advances in Complex Systems*, 11(4), 529–549.
- HUET, S. & Deffuant, G. (2010). Openness Leads To Opinion Stability and Narrowness To Volatility. *Advances in Complex Systems*, 13(3), 405–423.
- ISENBERG, D. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, 50(6), 1141–1151.
- JAGER, W. & Amblard, F. (2005). Uniformity, Bipolarization and Pluriformity Captured as Generic Stylized Behavior with an Agent-Based Simulation Model of Attitude Change. *Computational and Mathematical Organization Theory*, 10, 295–303.
- KELLY, C. (1988). Intergroup differentiation in a political context. *British Journal of Social Psychology*, 27(4), 319–332.
- KURMYSHEV, E., Juárez, H. A., & González-Silva, R. A. (2011). Dynamics of bounded confidence opinion in heterogeneous social networks: Concord against partial antagonism. *Physica A*, 390, 2945–2955.
- LAZARFELD, P. F., & Merton, R. K. (1954). Friendship as a Social Process: A Substantive and Methodological Analysis. In

- Berger, M., Abel, T., & Page, C. H. (Eds), *Freedom and Control in Modern Society*. New York: Van Nostrand. Leenders.
- LEONARDELLI, G. J., Pickett, C. L., & Brewer, M. B. (2010). Optimal distinctiveness theory: A framework for social identity, social cognition, and intergroup relations. *Advances In Experimental Social Psychology*, 43, 63–113.
- LORENZ, J. (2007). Continuous opinion dynamics under bounded confidence: a survey. *International Journal of Modern Physics C*, 18(12), 1819–1838.
- LORENZ, J. (2010). Heterogeneous bounds of confidence: Meet, discuss and find consensus! *Complexity*, 15(4), 43–52.
- LUM, C. & Kennedy, L.W. (Eds.) (2012). *Evidence-Based Counterterrorism Policy*. New York and Heidelberg: Springer.
- MÄS, M., Flache, A. & Helbing, D. (2010). Individualization as driving force of clustering phenomena in humans. *PLoS Comput. Biol.* 6(10), e1000959.
- MCCAULEY, C. & Mosakalenko, S. (2005). Mechanisms of Political Radicalization. *Terrorism and Political Violence* 17(3), 415-433.
- 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. <http://jasss.soc.surrey.ac.uk/9/1/13.html>
- SHERIF, M. & Hovland, C.I. (1961). *Social Judgment*. New Haven, CT: Yale University Press.
- SHERIF, M. (1966). *In common predicament: Social psychology of intergroup conflict and cooperation* Boston, MA: Houghton Mifflin.
- TAJFEL, H. (1974). Social identity and intergroup behavior. *Social Science Information*, 13, 65-93.
- TAJFEL, H. (1975). The exit of social mobility and the voice of social change. *Social Science Information*, 14, 101-118.
- TAJFEL, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. London: Academic Press.
- TAJFEL, H. & Turner, J. C. (1986). The social identity theory of inter-group behavior. In Worchel, S., & Austin, L. W. (Eds.), *Psychology of Intergroup Relations*. Chicago: Nelson-Hall.
- TURNER, J. C. (1975). Social comparison and social identity: Some prospects for intergroup behavior. *European Journal of Social Psychology*, 5, 5-34.
- TURNER, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D. & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Blackwell, Oxford.
- URBIG, D. & Malitz, R. (2007). Drifting to more extreme but balanced attitudes: Multidimensionnal attitudes and selective exposure. *4th Conference of the European Social Simulation Association*, 10-14 September, Toulouse.
- WEISBUCH, G. (2004). Bounded confidence and social networks. *The European Physical Journal B – Condensed Matter and Complex Systems*, 38(2), 339–343.
- WOOD, W., Pool, G. J., Leck, K. & Purvis, D. (1996). Self-definition, defensive processing, and influence: the normative impact of majority and minority groups. *Journal of Personality and Social Psychology*, 71(6), 1181–1193.