

Laurent Salzarulo (2006)

A Continuous Opinion Dynamics Model Based on the Principle of Meta-Contrast

Journal of Artificial Societies and Social Simulation vol. 9, no. 1 <http://jasss.soc.surrey.ac.uk/9/1/13.html>

For information about citing this article, click here

Received: 29-Jul-2005 Accepted: 27-Oct-2005 Published: 31-Jan-2006



Abstract

We propose a new continuous opinion dynamics model inspired by social psychology. It is based on a central assumption of self-categorization theory called principle of meta-contrast. We study the behaviour of the model for several network interactions and show that, in particular, consensus, polarization or extremism are possible outcomes, even without explicit introduction of extremist agents. The model is compared to other existing opinion dynamics models.

Keywords:

Opinion Dynamics, Self-Categorization Theory, Consensus, Polarization, Extremism

Introduction

1.1

Self-categorization theory is a relatively new paradigm in social psychology aiming at "[explaining] the psychological emergence of the group-level properties of social behaviour" (<u>Oakes et al. 1994</u>, p. 101). In Salzarulo (<u>2004</u>), a formal model of self-categorization theory is proposed, and it is suggested that it could be used to build a new opinion dynamics model, which would be social-psychologically founded. As the scope of the theory of self-categorization is group processes, it deals fundamentally with situations where a great number of individuals interact. These situations typically generate complex collective phenomena, which are difficult to anticipate on the basis of the behaviour of individuals. Simulation is a reliable way of exploring the collective dynamics resulting from the hypotheses made on the individual level. In this paper, we investigate the behaviour of this opinion dynamics model, which we call meta-contrast (MC) model. We will compare the results of the simulation to what is expected according to the theory, and to other opinion dynamics models.

1.2

We will first summarize some important aspects of self-categorization theory and describe other existing opinion dynamics models. Section $\underline{2}$ is devoted to the presentation of the MC model and is followed, in section $\underline{3}$, by the simulation results and a discussion.

1.3

Experiment proved that people define themselves and behave differently in different situations. The same person may, for example, think and act typically as a business-man during working time (being serious, wearing a suit, ...) but behave like a typical supporter during football matches (shouting, drinking, ...). Self-categorization theory (SCT, Turner et al. 1987; Oakes et <u>al. 1991; Oakes et al. 1994</u>) makes the hypothesis that one's identity is defined by the set of individuals with whom one identifies. Such a set is called a self-category, or ingroup. Other social categories are built to identify other people of the context; they are called outgroups. Identity thus depends on the present social context of an individual. SCT aims at predicting which social categories are going to form in a given social context, and at finding the prototypical behaviour of these categories. The idea is that we group together similar individuals to form a social category, and that we separate different individuals into different categories. However, there is no absolute measure of how similar two individuals are. Individuals can only be judged as similar relatively to other individuals in the context. For example, two individuals speaking French will not feel very similar if they cross in Paris, whereas they will feel very close to each other if they cross in a small village of the Amazonian forest. In accordance with what is called the principle of meta-contrast, SCT predicts that a given set of individuals will be more likely to be perceived as a category if the mean difference between this set of individuals and all other individuals of the context is perceived as larger than the mean difference between the individuals within this set.

1.4

SCT also postulates that the behaviour or the opinion we adopt is chosen in such a way as to be prototypical of the self-category. "The most prototypical response (argument, position, member, etc.) – the one that best represents and exemplifies the agreement of ingroup members - will tend to be perceived as most correct and be most valued: it will embody the (most) normative response in a given context." (Turner et al. 1987, p. 81). Thus, to identify with a group is to try to resemble as much as possible what we think is prototypical of this group. According to the principle of meta-contrast, the prototypicality of an individual (for his/her category) grows to the extent that the mean difference between this individual and members of the other groups is large compared to the mean difference between this individual and the other members of his/her group. The prototype of a category doesn't only depend on the category considered but also on the whole context, since the inter-category differences are taken into account in the choice of the prototype. Hence, not only the categories but also the prototypes are context-dependent. The fact that the inter-category differences intervene in the choice of the prototype also implies that it is not necessarily centred in its group: the prototype has to be similar to the members of its group, but also dissimilar from the individuals of other groups.

1.5

Group polarization is usually defined as the tendency of the average response of group members on some dimension to become more extreme towards the initially preferred pole after group discussion than the average of their initial individual responses. Conformity to the prototype can lead to convergence on the mean when the prototype is close the the initial average opinion as well as to polarization when the context is such that the prototype is off-centred. Hence, in the frame of this theory, polarization is an intra-group process (McGarty et al. 1992). The presence of extremists is not necessary to induce group polarization. To the contrary, mere exposure to a central tendency is sufficient (Isenberg 1986). Moreover, extremists can even lead to depolarization if they are categorized as outgroup.

Existing opinion dynamics models

1.6

Most of the existing opinion dynamics models (e.g. bounded confidence model (<u>Krause 2000</u>; <u>Deffuant et al. 2000</u>; <u>Dittmer 2001</u>; <u>Hegselmann and Krause 2002</u>), relative agreement model (<u>Deffuant et al. 2002</u>; <u>Amblard and Deffuant 2004</u>)) are based on an "attractive force" between similar opinions (whatever `similar' is in the corresponding model), but do not take into account any "repulsive force" between different social groups, which should exist according to SCT. As a result, these models are unable to foresee situations where the final opinions are more extreme than the most extreme initial opinion. In the bounded confidence (BC) model, each agent *i* has a

confidence level or uncertainty u_i , and considers the set l_i of agents whose opinion doesn't differ by more than u_i from his own. Formally, $l_i = \{j \mid |x_j - x_i| = u_i\}$. The opinion x_i of agent *i* is updated by averaging opinions of agents belonging to this set:

$$x_{i} \leftarrow |I_{i}|^{-1} \sum_{j \in I_{i}} x_{j} . \tag{1}$$

In other words, agents take into account opinions in an interval $[-u_i; +u_i]$ around their own opinion. By introducing asymmetric or opinion-dependent confidence intervals, agents can be made more `opened' to extreme opinions and convergence to extremism or polarization is observable.

1.7

In the case of pair interaction, a variant of the BC model can be defined by the following rule:

if
$$|x_i - x_j| < u_i, \quad x_i \leftarrow x_i + \mu (x_j - x_i)$$
, (2)

where μ is a parameter of the model controlling the speed of the dynamics (the BC model corresponds to the case $\mu = \frac{1}{2}$). To further explore the emergence of extremism, new models have been developed by adding a rule to update uncertainty in a similar way (Weisbuch et al. 2005). In particular, the relative agreement (RA) model (Deffuant et al. 2002; Amblard and Deffuant 2004) is defined in the following way. We consider the overlap h_{ij} of the two opinion segments $[x_i - u_i; x_i + u_i]$ and $[x_j - u_j; x_j + u_j]$, which is given by $h_{ij} = \min(x_i + u_i; x_j + u_j) - \max(x_i - u_i; x_j - u_j)$. The hypothesis is that the bigger the relative overlap between the two segments ($h_{ij}/u_i - 1$), the stronger the interaction. If $h_{ij} > u_i$, the following interaction takes place:

$$x_{i} \leftarrow x_{i} + \mu \left(\frac{h_{ij}}{u_{i}} - 1 \right) (x_{j} - x_{i}), \qquad (3)$$

$$u_{i} \leftarrow u_{i} + \mu \left(\frac{h_{ij}}{u_{i}} - 1 \right) (u_{j} - u_{i}) .$$

$$\tag{4}$$

Introducing "extremists" as agents with an extreme opinion and with a low uncertainty allows for large extreme clusters to emerge (single extreme convergence or bipolarization), even with a small initial proportion of extremists.

1.8

Experiments showed that extremists are indeed more self-confident, but from the point of view of SCT this is a result rather than a cause of differentiation: extremists have lower uncertainties and are more influential because they are more prototypical of an extreme category than a centrist of a central category. This has been shown to result from the principle of meta-contrast (Haslam and Turner 1995). There is no reason to suppose that extremists have *a priori* lower uncertainties. Moreover, SCT states that the cognitive process leading to extremism or polarization is the same as the one leading to centrism or consensus. In both cases, it can be explained in terms of conformity to the prototypical opinion, which in turn can be centred or not, depending on the situation. In particular, the presence of extremists before an interaction is not necessary to produce extremism. Hence, SCT would favour models (such as the following one and, as we will see, the model presented in this paper) in which the presence of extremists is a product of the interaction rather than the cause of extremism.

1.9

A model introducing a repulsive effect between two disagreeing agents was developed by Jager and Amblard (2005). It is inspired by social judgement theory and hereafter referred to as social judgement (SJ) model. Here, an agent *i* is characterized by his opinion x_i , a latitude of acceptance u_i and a latitude of rejection t_i , with $t_i > u_i$. The idea is that an agent meeting another agent having similar opinion (within the range $\pm u_i$) will try to agree with him, and that if the other agent has a very different opinion (outside of the range $\pm t_i$), he will reject his opinion and shift away from it. Formally, when two agents having opinions x_i and x_j meet, opinion x_i is updated according to:

$$if |x_i - x_j| < u_i, \quad x_i \leftarrow x_i + \mu (x_j - x_i) \quad \text{(assimilation rule)},$$

$$if |x_i - x_j| > t_i, \quad x_i \leftarrow x_i + \mu (x_i - x_j) \quad \text{(contrast rule)}.$$

$$(6)$$

Iterating these rules on a homogeneous population may lead to different kinds of behaviour, depending on the values of u_i and t_i . In particular, there can exist different number of opinion clusters. If one cluster remains (consensus), its opinion is close to the initial average opinion. If more than one cluster emerge, then some of them may shift to the extreme opinions, which captures a polarization phenomenon.

1.10

The models of this section will be compared to the one presented in this paper. We will see that our model produces results that are very similar to those obtained by the SJ model.

The meta-contrast model

2.1

This section is devoted to the presentation of the formal model derived from the principle of meta-contrast of SCT (see <u>Salzarulo 2004</u> for more details). Let's suppose an agent knows the opinions of *n* individuals on a given topic (which could be a political position, the extent to which they feel patriotic, or whatever opinion which is salient in the situation), and let's call $x_i \in$

[0;1] the initial opinion of individual *i* regarding this topic. We define $X = (x_i)_{i=1}^n$ as the

context, that is the set of opinions of all individuals. For the same `objective' situation, X can theoretically vary from individual to individual, according to one's knowledge of the situation.

We are looking for opinions that are highly prototypical of a certain category among all $x \in [0;1]$. According to the principle of meta-contrast, opinion x tends to be more prototypical of a certain social group to the extent that it maximizes its mean distance with opinions of the outgroup members compared to its mean distance with opinions of the ingroup members. We first suppose that individuals whose prototypical opinion is x are individuals whose opinion is close to x. This closeness is measured by a fuzzy membership function $\mu(x, x_i)$ determining how individual with opinion x_i a priori belongs to the category having the prototypical opinion x (this is only a first approximation and doesn't coincide with the final membership given by the model). $\mu(x, x_i)$ should be a decreasing function of $|x - x_i|$, but the model is not very sensitive

to the precise shape of
$$\mu$$
. In order that $\mu(x, x) = 1$ and $\lim_{|x-x_i|\to\infty} \mu(x, x_i) = 0$, we choose

$$\mu(x, x_{\rm i}) = \exp\left(-\frac{(x-x_i)^2}{w^2}\right), \qquad (7)$$

where $w \in [0;1]$ is a parameter of the model which can be interpreted as a typical group width (in the opinion space). By choosing such a function, we assume that agents are not aware of being there *as* group members, and that they don't know *a priori* others' group membership.¹

2.2

The intra-category distance (i.e. the mean distance between opinion x and the opinions of other ingroup members) is then given by

$$d_{\text{intra}}(x, X) = \frac{\sum_{i=1}^{n} \mu(x, x_i) \ (x - x_i)^2}{\sum_{i=1}^{n} \mu(x, x_i)} .$$
(8)

The inter-category distance (i.e. the mean distance between opinion x and the opinions of all outgroup members) is

$$d_{\text{inter}}(x, X) = \frac{\sum_{i=1}^{n} (1 - \mu(x, x_i)) (x - x_i)^2}{\sum_{i=1}^{n} (1 - \mu(x, x_i))} .$$
(9)

A measure of how opinion x is prototypical of some social group is given by how large its intercategory distance is, compared to its intra-category distance. We compare these two distances using a weighted difference that defines the prototypicality function:

$$P(x, X) = a \cdot d_{inter}(x, X) - (1 - a) \cdot d_{intra}(x, X) , \qquad (10)$$

where $a \in [0;1]$ is a new parameter called outgroup aversion. If a = 0, the maxima of *P* are located in the centre of the groups, i.e. in areas with a high density of opinions. In this case, there is no repulsive effect due to the outgroups. a > 0 means that each maximum of *P* will be close to one group centre, but slightly repulsed by every outgroup. If a = 1 (in fact as soon as a > 0.5), *P* gets higher and higher as *x* goes away from any group. A local maximum of *P* represents an opinion that will be recognized as prototypical by people having a similar opinion. Hence, a category is formed around this prototype by all individuals whose opinion is closer to that prototype than to all other maxima of *P*. The main purpose of the MC model was to predict the categorization that would occur within a given context (an applet performing categorization according to this model can be found at

http://wwpeople.unil.ch/laurent.salzarulo/en/categorize/), but there are two ways of turning it into an opinion dynamics model following SCT. The first one is given by the prediction that, provided ingroup identification is sufficient, individuals tend to adopt the prototypical opinion of their ingroup (Turner et al. 1987, chap. 4). This phenomenon has often been observed experimentally (e.g. Haslam and Turner 1995) and the opinion shift caused by prototype conformity has been shown to be able to last in the long term (e.g. David and Turner 1996, 1999). It is the approach we will adopt in paragraphs 3.14 and following. The second way is more indirect. It is predicted that the individual holding an opinion which is the closest to the ingroup prototypical position, but this individual feels no pressure to move away from it. "Conformity to the stereotypical ingroup position, therefore, implies convergence on the most prototypical member" (Turner et al. 1987, p. 81). For our opinion dynamics model, this means that agents adopt the opinion of the agent who has the most prototypical opinion for the ingroup. In summary, the whole updating rule for agent *i* having opinion x_i is:

1)

given X as the set of opinions agent *i* perceives in his neighbourhood, compute P(x, X);

2)

find the position x^* of the local maximum of *P* which is the closest to x_i (this is the prototypical opinion for *i*);

3)

find in *i*'s neighbourhood which agent has the opinion which is the closest to x^* . This agent has the most prototypical opinion of *i*'s category;

4)

adopt the opinion of the most prototypical neighbour.

The advantage of adopting the opinion of the most prototypical agent instead of directly adopting the most prototypical opinion is that no opinion is created during interaction. Hence, all opinions stay in the initial [0;1] interval, which must be ensured if we iterate the process in order to reach a steady state. This is why we use this update process for our first simulations (paragraphs 3.3 and following)

2.3

We may imagine that individuals have opinions about several topics simultaneously. In this case, the model can be generalized straightforwardly into a multidimensional one by replacing $(x - x_i)^2$ by $||x - x_i||^2 = \sum_{\alpha=1}^{D} (x_{\alpha} - x_{i\alpha})^2$ in (7), (8) and (9), where *D* is the number of different

topics, and $x, x_i \in [0;1]^D$. This would allow to study the simultaneous dynamics of several opinions per individual, but it is not the purpose of this paper.

2.4

In Salzarulo (2004), parameters *a* and *w* were adjusted to fit a set of experimental data obtained by Haslam and Turner (1995). Assuming that outgroup prototypes are perceived the same way as ingroup prototypes, the calibrated values were a = 0.08 and w = 0.36. It was then showed that it could reproduce known experimental phenomena addressed by social-psychology, such as group polarization and recategorization in a new context.

2.5

To show a bit more intuitively how two agents interact through this model, let's consider their opinions x_1 and x_2 . We will see where the prototypical opinion of agent 2 is in function of his relative position with agent 1. Figure <u>1</u> shows the prototypicality curve obtained for $x_1 = 0.3$ and $x_2 = 0.7$ (we maintain *a* and *w* to their calibrated values for this example). In this case, the local maximum of P(x, X) lying close to x_1 – that is the prototypical opinion of x_1 – is farther from x_2 than x_1 is. The new opinion of agent 1, then, will get more extreme after interaction with agent 2. This is an important feature of this model, that distinguishes it in particular from the BC and RA models. While the updating rule of other models is essentially an averaging process among similar opinions, this model includes a kind of repulsive force between two opposing opinions, allowing opinions that are more extreme than the existing ones to emerge during interaction.

2.6

In the situation where x_1 and x_2 are closer to each other than previously, the prototypicality curve has a single maximum located halfway between x_1 and x_2 for the same values of both parameters *a* and *w* (figure <u>2</u>).

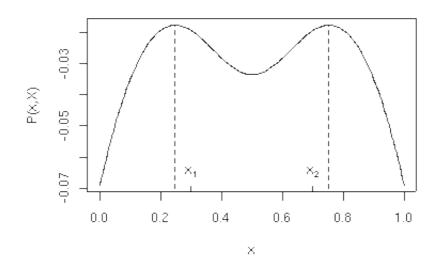


Figure 1: Prototypicality curve P(x, X) for the context $X = (x_1 = 0.3, x_2 = 0.7)$. a = 0.08, w = 0.36.

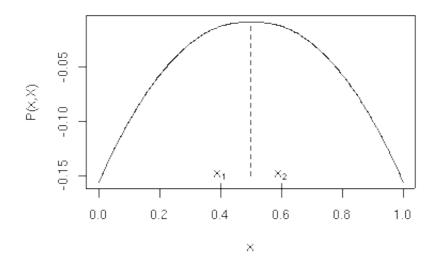


Figure 2: Prototypicality curve P(x, X) for the context $X = (x_1 = 0.4, x_2 = 0.6)$. a = 0.08, w = 0.36.

Here, the new opinion of agent 1 will be closer to that of agent 2 than it was before, and the process will eventually lead to consensus. In order to see systematically what happens, we measure the value of the first derivative of P at x_1 as a function of the relative position of x_1 and x_2 , $(x_1 - x_2)$. This can be considered as a force pushing the prototypical opinion to the right if it is positive, and to the left if it is negative. As we would like to study the direction of the force relatively to x_2 , we choose the sign of the derivative so that it is positive if the force is attractive (the new opinion of agent 1 will be closer to x_2 than it previously was), and negative if the force is repulsive (the new opinion of agent 1 will be farther from x_2 than it previously was). The result is shown on figure <u>3</u>.

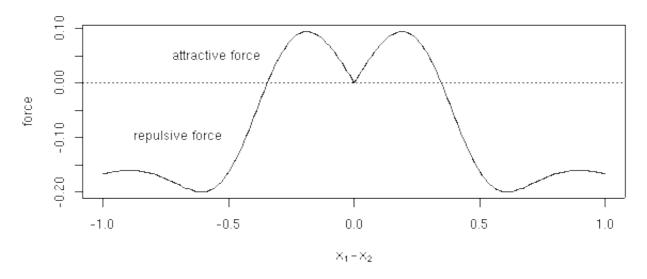


Figure 3: Force attracting or repulsing agent 1's opinion as a function of his relative position with agent 2, $(x_1 - x_2)$. The intensity of the force is given by $\frac{\partial P}{\partial x}(x_1)$.

This function has the shape of a Stetson hat. When $x_1 = x_2$, both agents already share the same opinion, which is also the prototypical one. Both opinions will remain the same, and there is no force exerting on them. If the distance between the two agents increases but remains small, they still find themselves similar (the prototypicality function keeps a single maximum located on the shared prototypical position) and their opinions have a tendency to get closer and closer to each other. Here, there is an attracting force between the two opinions. As the distance

increases again, the force vanishes and then becomes repulsive. There is a point (which depends on both parameters *a* and *w*) where the initial difference between the agents is too big to allow them to come to a consensus. Each agent is considered as outgroup by the other agent (two maxima have appeared in the prototypicality function, located on two distinct prototypical positions), and their opinions extremize. This behaviour is in accordance with SCT and experimental facts. This example shows the behaviour of the model in a particular situation where only two agents meet. Adding some more agents to the context may totally change the shape of the prototypicality curve. As shown in Salzarulo (2004), two agents categorizing themselves separately and therefore showing polarized prototypical opinions can categorize themselves together and have consensual opinions if they are put in the presence of a third agent which is much more different from themselves than they are. This phenomenon is well known in social psychology (Oakes et al. 1994). Thus the behaviour of the model is highly context-dependent and difficult (if not impossible) to anticipate analytically in situations involving many agents. This illustrates well the utility of simulation.

2.7

The most important difference between BC or SJ models and the MC model may precisely be that similarity between agents – that is the most different opinion with which an agent would tend to agree – is context-dependent in the latter model but absolute in the former ones. We will see how this difference affects the collective behaviour of these models.

Results and discussion

3.1

In this section, we investigate the collective outcomes of the self-categorization process. In order to make our results comparable to those obtained by other models, we first use random sequential update, which is the updating process used in the study of the RA and SJ models. For the sake of simplicity, we start our simulations (paragraph 3.3) by supposing that agents can communicate with all other agents in the population. Then, we introduce in paragraph 3.8 a small-world network to model a more realistic social network among agents. In this section, we compare the behaviour of the system to the one obtained in the full mixing condition, and study the effects of the randomness of the network. Finally, we argue that parallel updating is a more appropriate dynamics to model self-categorization (it is also used in Hegselmann and Krause (2002) to study the BC model), and thus study this interaction dynamics in paragraph 3.14 and following.

3.2

In section 2 we said that parameter $a \in [0;1]$. Actually, if a > 0.5, the repulsive effect of outgroups dominates over the attraction of the ingroup such that prototypical opinions are always located at $-\infty$ or $+\infty$ for all groups. Moreover, there is essentially no change in the dynamics when a varies from 0.3 to 0.5. Therefore, in the following we restrict the study of the model to the region $a \in [0;0.3]$. The code used to make all simulations is <u>here</u> (it is written in C++ and was compiled on a Linux system).

Random sequential update in a full mixing condition

3.3

We study here a population of 100 agents. At each time step, an agent is randomly chosen and updated applying the self-categorization rule given above. Agents are fully connected, which means that every agent exactly knows the opinion of any other agent in the population. Opinions are initially uniformly distributed between 0 and 1 in the population. In this situation, we observe that agents will tend to form one or several clusters sharing the same opinion. We first study the effects of varying parameters *a* and *w* by measuring the final number of different opinions remaining in the population after the system reaches a steady state. Two opinions are considered as equal when their absolute difference is smaller than 10^{-4} (it is usually the case that two such opinions are exactly equal – remember opinions are only copied and spread in the population, there is no creation or modification of existing opinions). Figure <u>4</u> shows the final

number of opinion clusters for different values of *a* and *w*. Each single value is obtained by averaging results over 100 runs. Each run is composed of 1500 iterations. Hence, each agent is updated 15 times on average.

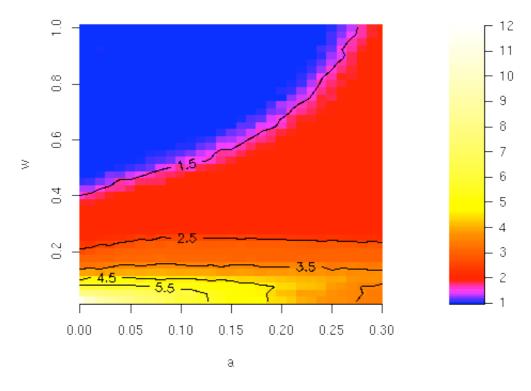


Figure 4: Final number of opinion clusters with random sequential update of 100 fully connected agents. Results are averaged on 100 runs of 1500 iterations each.

The upper left (blue) part of the diagram is the consensus region of the parameters space where only one opinion cluster forms: all agents share the same opinion. It corresponds to high values of *w* and small values of *a*, that is a situation where very different opinions are still considered as ingroup opinions and where aversion for the outgroups is low. A second big (red) area, at the centre of the diagram, corresponds to the formation of two clusters. In this case, the two clusters almost always have extreme opinions (see below), and thus we can speak about bipolarization: modulo statistical fluctuations, agents that have an initial opinion below 0.5 finish with an opinion of 0, while other agents finish with an opinion of 1. Below this central region, the smaller *w* and *a*, the more the opinion clusters get numerous. Intuitively, what happens here is that agents change less and less their opinion for another, and the situation remains close to the initial condition. Indeed, small *w* means that very few other opinions are considered as ingroup opinions, and small *a* means that there is few outgroup aversion and thus no reason to polarize.

3.4

We can compare this phase diagram with some results obtained analytically. It is possible to write an algebraical expression for *P* in the case of an exact uniform distribution of opinions (see <u>appendix</u>). In the full mixing condition, all opinions are known by each agent and thus the prototypicality function is the same for each of them. What differs from agent to agent is the very local maximum he will choose as his particular most prototypical opinion (that is the maximum of *P* the closest to his own opinion). For an exact uniform distribution, the prototypicality curve has either one (central) or two maxima, depending on the values of *a* and *w*. If *P* has a single maximum at 0.5, then the whole population will adopt this opinion and consensus is reached after the first iteration. If *P* has two maxima, say at x_1 and x_2 with $x_1 < x_2$, they are located symmetrically around 0.5. Thus, the half of the population having opinion below 0.5 will adopt opinion x_1 , and the other half will adopt opinion x_2 . In this case, the new prototypicality curve can be computed. After this iteration, either the two clusters merge, or their opinion difference is bigger than previously. In the former case, population has reached consensus and in the latter, it evolves toward bipolarization. This process leads to the diagram

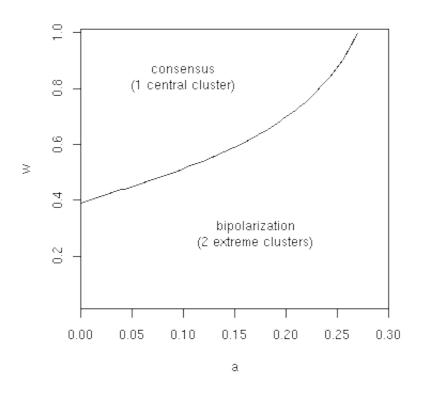


Figure 5: Final number of opinion clusters with random sequential update. Theoretical curve obtained for an exact uniform distribution of opinions.

Things are much simpler in this case. The curve on the figure separates the two only theoretically possible phases of the system. Above the curve (high *w*, small *a*), consensus is reached exactly on the average opinion. Below the curve, half of the agents (having opinion below 0.5) adopt opinion 0 while the other half adopt opinion 1. We may first note that this curve is in excellent agreement with the border (corresponding to the 1.5–level curve) separating the consensus and bipolarization phases obtained through simulation. What is surprising is that there is no theoretical possibility of obtaining a stable multiple (i.e. more than two) clusters final situation. This is possible in the simulation because the initial distribution of opinion is not perfectly uniform. A higher local density of opinion around this maximum will see it as their prototypical opinion and adopt it, making this maximum even higher as the local density increases. In summary, the system is very sensitive to the initial distribution of opinions and statistical fluctuations around the uniform distribution are responsible for a much richer behaviour. This illustrates well the utility of simulation in such a case.

3.5

In order to study the transition between the different phases, we plot on figure $\underline{6}$ a cut in the diagram of figure $\underline{4}$ for w = 0.6. The curves represent the final number of opinion clusters as a function of *a* for different sizes of the population, ranging from 25 to 900 agents.

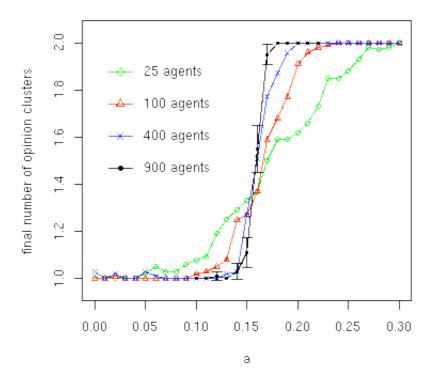


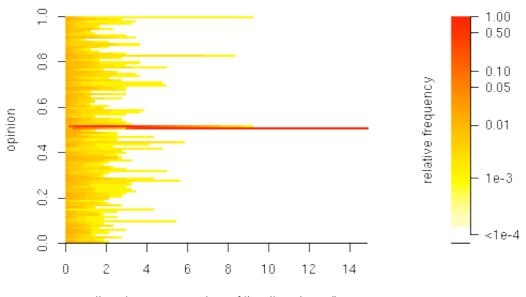
Figure 6: Final number of opinion clusters with random sequential update for different numbers of agents in the simulation (w = 0.6). Results are averaged on 100 runs. The number of iterations in each run is 15 times the number of agents. Vertical bars represent the 95% confidence interval for the 900 agents condition.

As in figure 4, we can see that consensus arises when $a \lesssim 0.15$ and bipolarization when $a \gtrsim$

0.15. The transition gets sharper and sharper as the population size increases. Moreover, the variance of the number of clusters over different runs vanishes for very small or very high *a*, but is important in the transition region (see the confidence intervals on the graph). The same kind of behaviour can be observed between the 2-clusters region and the 3-clusters region, and so on. This is a clue that the system is undergoing a first order phase transition between these different regions.

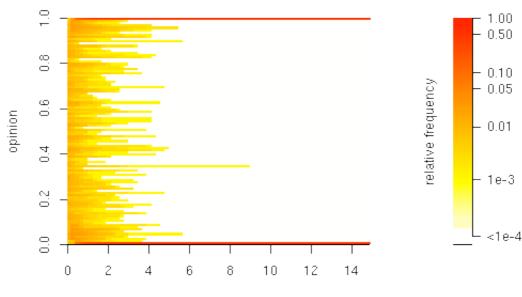
3.6

The distribution of the opinions in a population of 900 agents as a function of time is plotted for four individual runs in figures $\frac{7}{10}$ to $\frac{10}{10}$.



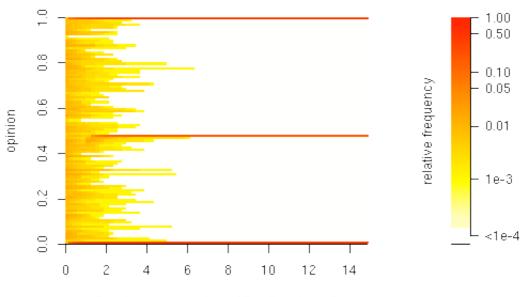
time (average number of iterations/agent)

Figure 7: Opinion distribution as a function of time with random sequential update on a population of 900 fully connected agents. Individual run for a = 0.05 and w = 0.8. Colour codes for the height of the opinion histogram at a given time.



time (average number of iterations/agent)

Figure 8: Opinion distribution as a function of time with random sequential update on a population of 900 fully connected agents. Individual run for a = 0.15 and w = 0.4. Colour codes for the height of the opinion histogram at a given time.



time (average number of iterations/agent)

Figure 9: Opinion distribution as a function of time with random sequential update on a population of 900 fully connected agents. Individual run for a = 0.15 and w = 0.2. Colour codes for the height of the opinion histogram at a given time.

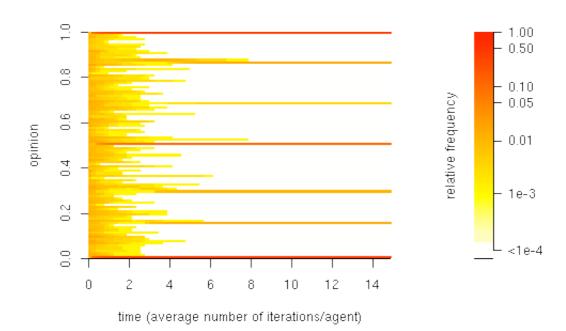


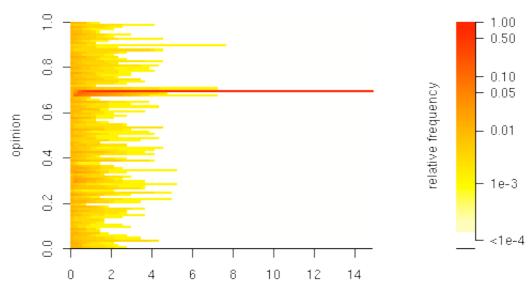
Figure 10: Opinion distribution as a function of time with random sequential update on a population of 900 fully connected agents. Individual run for a = 0.05 and w = 0.05. Colour codes for the height of the opinion histogram at a given time.

Figure <u>7</u> shows the evolution of the population in the consensus region. After each agent has been updated several times, the whole population has got the same opinion, which is close to the initial average (0.5). Moving to the bipolarization region (fig. <u>8</u>), one sees that, since the very beginning of the simulation, each time an agent is updated, it adopts one of the two extreme opinions. Unlike BC, RA and SJ models, there is no gentle shift of the opinions to their final values. The prototypical opinions here are immediately the extreme ones. Moreover, the formation of two clusters is possible in BC and RA models, but extremism is observed only if uncertainty is asymmetric or opinion dependent (BC), or if extremist agents are introduced explicitly with initial extreme opinion and low uncertainty (RA). In this simulation, the initial

distribution of opinions is uniform and, apart from their initial opinion, all agents are identical. Hence, nothing allows to foresee the bipolarization of the population – extremism is emergent. In the third figure (9), three opinions remain in the population after a stationary state has been reached. Around 83% of the population adopts one of the extreme opinions, while the small part left adopts a central opinion. This shows that, in accordance with SCT, extreme or central opinions can be adopted using the same process. There is no need to introduce *a priori* personal differences between extremists and centrists. Finally, fig. <u>10</u> is an example where 7 opinions coexist. The two biggest clusters are the extreme ones (sharing 83% of the population). The central cluster is smaller (12%), and the rest of the population (5%) is divided into the four remaining clusters. Here again, there is no opinion drift; the remaining opinions are prototypical almost from the beginning on. The effect of time is only to make every other opinion disappear as each agents adopts one of the prototypical opinions. Let's remark that all these kinds of convergence are exactly the one obtained with the SJ model. A better comparison between these two models would then be suitable.

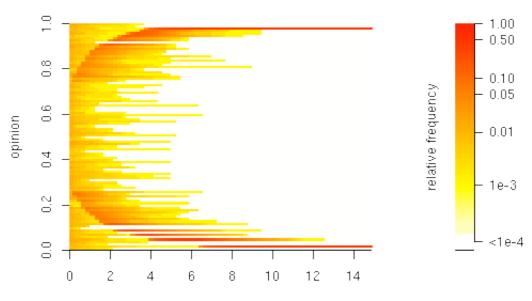
3.7

Two other individual runs are represented on figures 11 and 12.



time (average number of iterations/agent)

Figure 11: Opinion distribution as a function of time with random sequential update on a population of 900 fully connected agents. Individual run for a = 0.0 and w = 0.4. Colour codes for the height of the opinion histogram at a given time.



time (average number of iterations/agent)

Figure 12: Opinion distribution as a function of time with random sequential update on a population of 900 fully connected agents. Individual run for a = 0.0 and w = 0.4. Colour codes for the height of the opinion histogram at a given time. White horizontal stripes are caused by opinions that disappeared before they became prototypical.

Both runs were made exactly on the border between the consensus region and the bipolarization region (a = 0.0, w = 0.4). The first run converges to consensus but in the second one, the population splits into two clusters with opinions shifting to the extremes. Even at the boundary with the consensus region, when two clusters form, they become extreme. This shows that there is no intermediate state between consensus and extreme bipolarization (we shall see in paragraph 3.14 that this is not true anymore if the process is not iterated).

Random sequential update in a small-world network

3.8

Until now, we have supposed that agents were fully connected. In this section, we let agents interact through a 2D small-world network (SWN, <u>Watts and Strogatz 1998</u>) and see how this affects the model. To build our SWN, we start from a 50×50 square grid with periodic boundary conditions. One agent is placed on each node of this grid. We consider that each agent is linked with (i.e. knows the opinion of) himself and his eight immediately surrounding neighbours, and that communication is symmetric. This corresponds to a Moore neighbourhood of radius 1, and to a connectivity k = 8 in the SWN formalism. Then, each link of this network (except the ones linking an agent to himself) is rewired with probability p to another randomly chosen agent. Thus, p = 0 corresponds to a square lattice with Moore neighbourhood and p = 1 to a random network with average connectivity k. The two parameters k and p characterize the network. SWNs are known both to be highly clustered and to have a short average distance between two randomly chosen agents, which make them good models of real social networks.

3.9

Instead of studying, as before, the number of clusters (i.e. the number of different opinions in the population), we will here measure and study the variance of the opinions. Indeed, when the system is spatialized, several clusters can exist with almost the same opinion that are separated spatially and therefore cannot communicate. If they were connected, they would converge to a single opinion cluster, but the spatial segregation prevents them to do so. This could result, for example, in a situation where the population is composed of a lot of clusters with different opinions, but all of them close to 0, and another set of clusters with similar opinions close to 1. Intuitively, we would say that the opinions have polarized and that there remains roughly two main opinions, but counting the number of different opinions would give us a much bigger

number which wouldn't describe well the situation. Moreover, in the spatialized case, there can be stable patterns of gradient of opinions joining two different clusters (see snap-shots on figures <u>17</u> and <u>18</u>). Here again, it doesn't seem reasonable to count each different opinion of the gradient. Of course, it is impossible to fully describe the situation using a single number, but we found that the use of the variance, measuring the dispersion of the opinions in the population, captures the system's characteristics we want to observe.

3.10

Initially, the distribution of the opinions is uniform between 0 and 1, and the corresponding variance is 1/12. The maximal value of the variance is 1/4, it corresponds to an extreme bipolarization (half of the population has opinion 0 and the other half 1). Naturally, the lowest value is 0 and corresponds to consensus. In order to get an intuitive measure of the situation, we "normalize" the variance by dividing it by its value at the beginning of the simulation, that is 1/12. Hence, the normalized measure is smaller than one if opinions had a tendency to group together toward consensus, and bigger than one if opinions had a tendency to polarize. A value of 0 indicates a consensus and a value of 3 extreme bipolarization. Other values can be obtained by several different distribution, but as an example, a value of 2 is obtained if one third of the population has opinion 0, another third opinion 1/2 and the rest opinion 1.

3.11

The normalized opinion variance is represented as a function of w and a for different values of p in figures <u>13</u> to <u>15</u>.

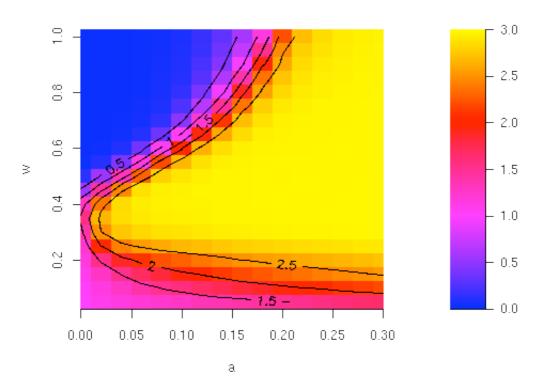


Figure 13: Normalized opinion variance for a population of 2500 agents in a small-world network with parameter p = 0 (square lattice with Moore neighbourhood of radius 1). Each point of the diagram is an average on 50 runs of 2.5 \cdot 10⁵ iterations each.

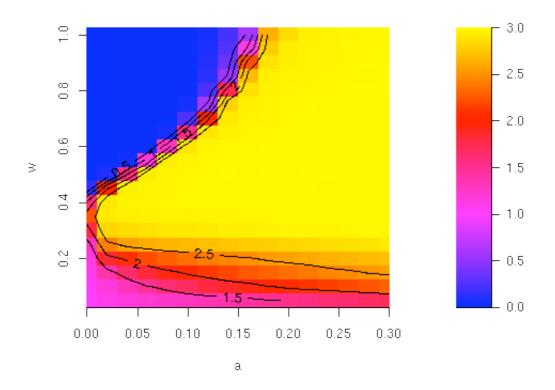


Figure 14: Normalized opinion variance for a population of 2500 agents in a small-world network with parameter p = 0.1. Each point of the diagram is an average on 50 runs of $2.5 \cdot 10^5$ iterations each.

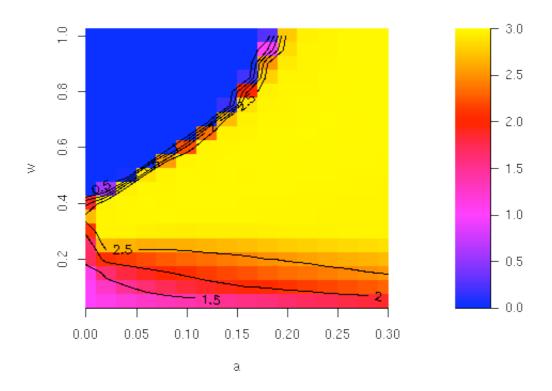


Figure 15: Normalized opinion variance for a population of 2500 agents in a small-world network with parameter p = 1 (random network). Each point of the diagram is an average on 50 runs of 2.5 \cdot 10⁵ iterations each.

First, we note that these phase spaces look very much like the one of the fully connected condition (fig. <u>4</u>). The upper left part of each diagram (in blue) is the consensus phase. Besides and below it, comes the bipolarization phase (we shall see that the transition between these two phases is not of the same type as the one of the full mixing case). In the bottom parts of these diagrams, for small *w* values, the normalized variance decreases again to a value of 1 (i.e. close

to its initial value) for the same reason as in the full mixing condition.

3.12

Even if it doesn't change the kind of convergence obtained, increasing the randomness p of the network has three effects on the behaviour of the system:

- 1. the pure consensus phase is enlarged toward bigger values of *a*;
- 2. the transition is sharper and sharper, as shown on figure <u>16</u>; intermediate states between consensus and polarization disappear;

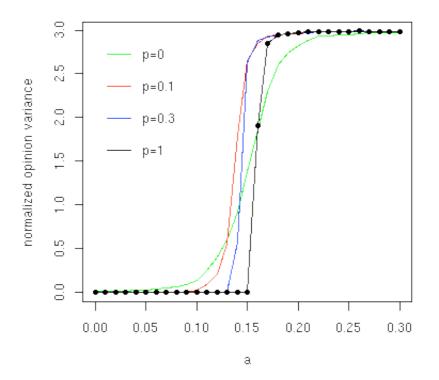


Figure 16: Normalized opinion variance for a population of 2500 agents in a small-world network (k = 8, p as indicated on the figure). These are cuts in the diagrams of figure 13, 14 and 15 for w = 0.8. Each point is an average on 50 runs of $2.5 \cdot 10^5$ iterations.

3. the convergence speed of the system increases. To measure the convergence speed, we do the following: each time an agent is updated, the absolute difference between his former and new opinions is stored. Then, at each iteration, we compute the average of these differences over the whole population and we call this mean last change of opinion. This measure is an overestimation of the opinion change in the population, since it may be that an agent has changed his opinion to his final one during his last update. Thus, the mean last change of opinion is 0 only if every agent has not changed his opinion during his last update. Noting that the mean last change of opinion decreases roughly exponentially with time after 5 iterations per agent, we can quantify the convergence speeds. If we adjust an exponential to fit these curves, then the opposite of the average number of iterations needed for the mean last change to decrease by a factor *e*). These convergence speeds are shown in table <u>1</u>.

Table 1: Convergence speed s obtained by supposing
an exponential decay of the mean last change of
opinion

s consensus bipolarization (a = 0.05, w = 0.8) (a = 0.15, w = 0.4)

k = 8, p = 0	0.07	0.31
k = 8, p = 0.1	0.06	0.64
k = 8, p = 0.3	0.12	0.99
k = 8, p = 1	0.28	0.88
full mixing	0.95	0.79

First, we note that, except for the full mixing condition, bipolarization converges much faster than consensus. When the extreme opinion propagates, then each time an agent is in contact with it, he adopts it and it is directly his final opinion, since no other opinion are more extreme than this one. On the other hand, agents change their opinion for a very long time in a consensus case until they find their final one, that has to take into account the influence of each group, even distant ones. Second, the higher *p*, the faster the convergence. As we have seen, small-world networks have short characteristic path lengths. Hence, as soon as p > 0, information propagates very quickly from one side of the network to another. In the bipolarization case, turning from p = 0 to p > 0 has a strong effect on the convergence speed, but a high or low value of *p* produces approximately the same outcome. Finally and unsurprisingly, the full mixing condition is in both cases among the fastest networks.

Increasing the randomness of the network can be interpreted in terms of a growing use of long distance communication channels (television, Internet, e-mail, ...) which provides a direct access to what's happening at the other side of the world. This reduces both the number of intermediates and the speed necessary to be informed of something. Our results suggest that in such a society people get more rapidly to a definitive opinion and that convergence ends either on clear consensus or clear extreme bipolarization, without intermediate state. Somehow paradoxically then, increasing long distance, immediate communication doesn't lead only to a more rapid consensus, but also to a more rapid polarization. Potential conflicts caused by initial divergence aren't avoided but further enhanced. Thus this model could partly explain how the "communication era" can simultaneously lead to the current rapid uniformization of culture (music, movies) in the world and to the emergence of extremists movements (populist parties in western countries, Islamic fundamentalism), supposing both phenomena really increased during the last decades.

3.13

We said before that the transition between the consensus and bipolarization phases is not a first order phase transition in the spatialized case. Indeed, for example for p = 0, we can see on figures <u>17</u> and <u>18</u> that in the transition region, there exists stable intermediate states between consensus and bipolarization. In these two examples, there are homogeneous clusters with extreme opinions, clusters with central opinions and smooth gradients between some of them.

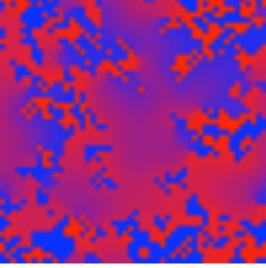


Figure 17: Final state of a population of 2500 agents on a 50×50 grid with Moore neighbourhood of radius 1 (after 2.5 \cdot 10⁵ iterations). Opinions

are represented by a colour ranging from blue to red, blue being opinion 0 and red opinion 1. a = 0.0 and w = 0.4.

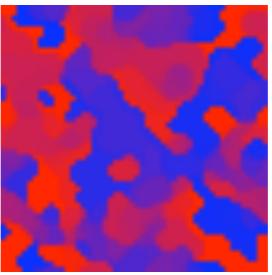
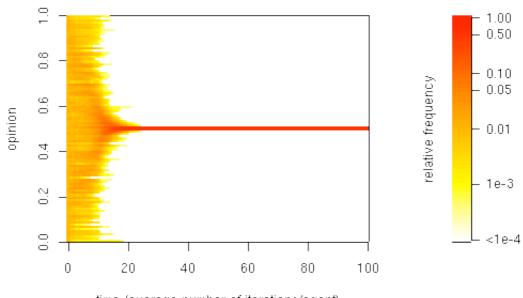


Figure 18: Final state of a population of 2500 agents on a 50×50 grid with Moore neighbourhood of radius 1 (after $2.5 \cdot 10^5$ iterations). Opinions are represented by a colour ranging from blue to red, blue being opinion 0 and red opinion 1. a = 0.18 and w= 1.0.

Moreover, in the transition region between consensus and bipolarization and for high values of p, convergence can be consensual, but to an opinion which is not necessarily central. Figures <u>19</u> to <u>21</u> show the opinion distribution versus time for three individual runs in the transition region (a = 0.0, w = 0.41).



time (average number of iterations/agent)

Figure 19: Opinion distribution as a function of time for a population of 2500 in a small-world network with p = 1 (a = 0.0, w = 0.41). Colour codes for the height of the opinion histogram at a given time.

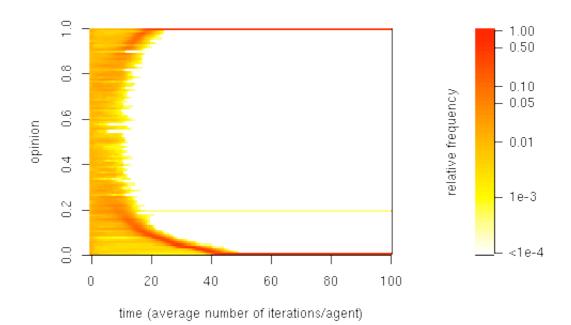


Figure 20: Opinion distribution as a function of time for a population of 2500 in a small-world network with p = 1 (a = 0.0, w = 0.41). Colour codes for the height of the opinion histogram at a given time.

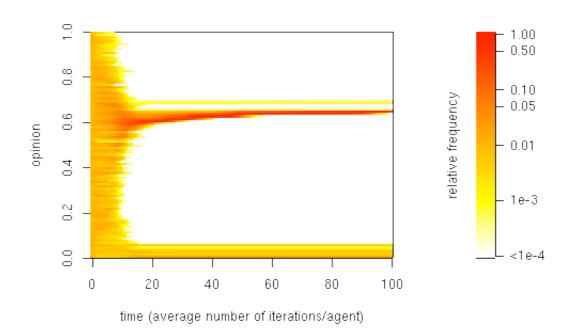


Figure 21: Opinion distribution as a function of time for a population of 2500 in a small-world network with p = 1 (a = 0.0, w = 0.41). Colour codes for the height of the opinion histogram at a given time.

For these parameter settings, we can observe consensus on the average opinion (fig. <u>19</u>), extreme bipolarization (fig. <u>20</u>) or, on figure <u>21</u>, a kind of convergence reminding the so called single extreme convergence observed in variants of the BC model and in the RA model (<u>Deffuant</u> <u>et al. 2002</u>). However, there are three main differences:

- 1. extremists or asymmetric uncertainties have to be entered explicitly in the BC and RA models, whereas all agents are identical in our simulation;
- during the experiment represented on figure <u>21</u>, the biggest cluster (96% of the population) has final opinion 0.65, while 4% (spatially isolated clusters) have an opinion close to 0. Since an outgroup is necessary to "push" the opinion of the big cluster toward high values, one deduces that the small isolated clusters are responsible for this shift.

Unlike in the BC and RA models, where agents are attracted by a small group of extremists, they are here repelled by a small group of opposite extremists (that emerged during interaction). This behaviour is similar to the one of the SJ model, but since it has not been tested on such a network, a direct comparison is impossible. The conditions required for this kind of convergence to appear could be compared with many hypotheses developed around the influence of minorities in group discussion (e.g. <u>Mackie 1987</u>; <u>Tindale et al. 1990</u>; <u>David and Turner 1996</u>, <u>1999</u>);

3. the final clusters obtained in the BC and RA models have opinions very close to the maximal value, hence they are really extreme. In the MC model, the final opinion is not extreme, but still very different from the average opinion. The final consensus in the case p = 0 is almost always reached on an opinion between 0.45 and 0.55, but values of 0.3 or 0.7 are still very frequent for p = 1.

Single parallel update of the opinions

3.14

We have studied the behaviour of the MC model when it is iterated until a stationary state is reached. This gave the possibility to compare our results with the ones obtained by other models using this kind of dynamics. Now, this is a problem when we think about the social meaning of our experiments: this dynamics means agents interact more than once on the same topic (we only considered the unidimensional version of the model) with the same agents. However, the principle of meta-contrast and SCT model in a single step the whole process of what happens between the beginning of a group interaction and its end. Hence, in our model, the time unit (an iteration) is well defined and corresponds to a whole group discussion, not just to a single pair interaction. To say we would repeat the same discussion on the same topic with the same participants in the same conditions until opinions stabilize would introduce strong biases, such as the memory of the agent reminding his previous position, his previous discussions, an *a priori* categorization of the other participants, boredom, etc. If we try to be more faithful to the theory we try to model, we have to study simulation results after the first iteration.

3.15

Moreover, the updating algorithm used until now states that agents adopt the most prototypical opinion found in their neighbourhood. This was used in order to ensure that opinions stay in the [0;1] interval during iterated interaction. According to SCT (see <u>introduction</u>), the opinion we tend to adopt is the one we think is prototypical of the ingroup. Conformity to the most prototypical ingroup member is seen as a consequence of this behaviour and therefore is not the primary behaviour we should implement. The prototypical opinion is rather constructed than observed, and doesn't necessarily exist in our surrounding. A more theoretically correct dynamics is to let agents adopt their really prototypical opinion instead of the most prototypical one in their neighbourhood. Thus we replace steps 3 and 4 of the updating algorithm by a single step:

3')

adopt opinion x^* .

This means that new opinions are created during interaction, in particular opinions that lie outside of the initial range [0;1].

3.16

We study a population of 100 fully connected agents whose initial opinions are drawn from a uniform distribution between 0 and 1. Since we want all agents to be updated during a single iteration representing a whole group discussion, we choose parallel updating (all agents are updated once simultaneously). The normalized opinion variance obtained in this case is shown in figure <u>22</u> for different values of parameters *a* and *w*.

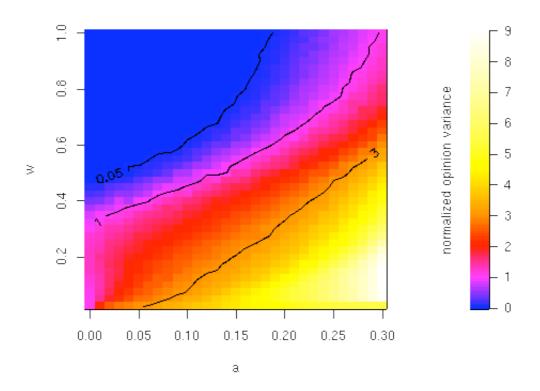


Figure 22: Simulation results for normalized opinion variance of 100 agents in a full mixing condition after a single iteration with parallel updating. Each point is an average over 100 runs.

Above the level curve 0.05, one obtains consensus, always very close to the initial average opinion. Between level curves 0.05 and 1, two clusters are observed, but they are very close to each other, since the normalized opinion variance stays below its initial value. Between level curves 1 and 3, we can speak about polarization, since opinions are more dispersed after interaction than before. The polarization gets even extreme (two clusters at 0 and 1) when we approach level curve 3. An interesting result is that the normalized variance can become higher than 3 in the bottom-right part of the diagram. In this case, polarization has generated opinions more extreme than the most extreme initial opinions, which can perfectly be observed experimentally (e.g. Haslam and Turner 1995) and should be a possible outcome of the model.

3.17

Theoretical results can be obtained using the prototypicality function computed for an exact uniform distribution (see <u>appendix</u>). A comparison between the theoretical curve and simulation results for w = 0.6 is shown in figure 23. The agreement between them is excellent: the average absolute difference over each point of the phase diagram is smaller than 0.15, for a measure that ranges from 0 to almost 9. This difference is mostly due to theoretical values that are higher than the empirical ones. A reason explaining this is that the actual initial distribution of opinions is not exactly uniform. It can be that the opinion density becomes locally high enough to induce a maximum in the prototypicality function and form a cluster, whereas this is impossible with the theoretical uniform distribution.

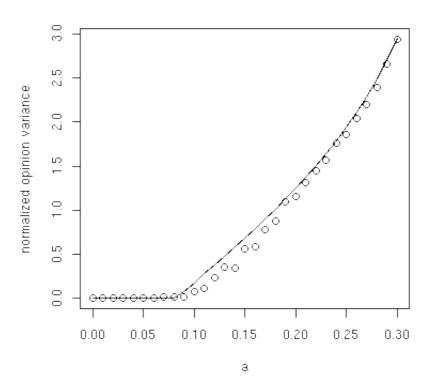


Figure 23: Cut in the diagram of figure <u>22</u> for w = 0.6. Dots (o) are simulation results and the curve is obtained numerically supposing an exact uniform distribution of opinions.

We see on figure 23 that the transition between the consensus phase and the region between level curves 1 and 3 is not of the same type as the one of the iterated model (compare with figures 6 and 16). Here, our order parameter (the normalized variance) suddenly becomes different from 0, but in a continuous way. Thus, this transition has more to do with a second order phase transition.

3.18

In the iterated version of the model (and for a sufficiently regular network), the possible outcomes were essentially either one central cluster or several clusters, two of which extreme. This is in accordance with the experimental fact that repeated attitude expressions and repeated exposure to a given stimulus, which corresponds to the iterated update, increases polarization (Brauer and Judd 1996). Yet if we look at the number of clusters in the single update situation (fig. 24), we can observe the formation of 3 or even 4 clusters, none of which are extreme.

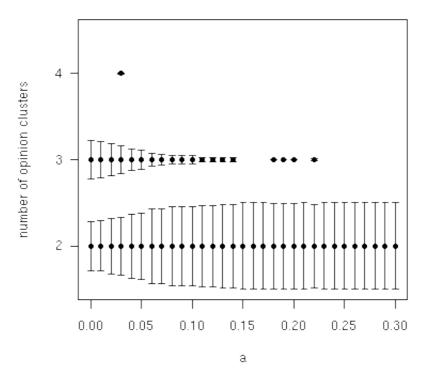


Figure 24: Number of opinion clusters in a population of 100 agents after a single parallel update for w = 0.2. The length of the vertical bars represents the proportion of experiments having the corresponding number of clusters. 100 experiments were made for each value of *a*.

Hence, with this update algorithm, consensus, bipolarization (and more generally extreme clusters) as well as simple clustering without polarization are observable.

Conclusion

4.1

The main features of the meta-contrast model are the following:

- when we iterate the self-categorization process in the full mixing condition, opinions group in one central cluster or several clusters, two of which are always extreme. As a result, the model generates consensus, bipolarization or multiple clusters depending on the values of both parameters *a* and *w*. All types of convergence are very similar to what has been obtained for the SJ model. It has been shown that either one or two clusters can be observed if we suppose an exact uniform distribution of the opinions. Fluctuations around this theoretical distribution are responsible for a much richer behaviour that could be exhibited through simulation;
- the different states reached by the system are essentially the same when agents interact through a small-world network, except that intermediate states appear between consensus and extreme bipolarization: population can form several homogeneous clusters spatially separated either by a sharp border or by a smooth opinion gradient. These intermediate states disappear as the randomness of the network increases. In this case, a small proportion of agents having extreme opinion can push the majority consensus away from the initial average opinion. Non-zero randomness in the small-world network, that is the ability to perform long distance communication provokes a rapid crystallization of the population not only into consensus but also to extreme bipolarization;
- when only one parallel iteration is done, virtually every outcome is possible: consensus, several clusters having extreme opinions or not. A significant part of the phase space represents runs in which the final opinions are more extreme than the most extreme

initial opinion.

The MC model was shown to be consistent with empirical facts at the individual and smallgroup level in Salzarulo (2004). The results obtained here at the collective level are coherent with the predictions of SCT. In particular, SCT argues that the same psychological process is at work when one adopts an extreme opinion or a central one. We could observe that small variations (random fluctuations) in the initial distribution of opinions could lead, for example, either to consensus or to extreme bipolarization. Moreover, this was obtained without changing the model's parameters if their value was well chosen. The initial presence of extremists with a particular behaviour wasn't necessary either. Thus, in accordance with the theory, a population of agents all having the same behaviour is likely to lead to several different collective outcomes. We showed further, and still in accordance with SCT and experimental facts, that final opinions in the population could be more extreme than the most extreme initial opinions.

4.2

The question of the similarity with the SJ model should be deepened. As we said in the introduction, the main difference between SJ and MC models is that the necessary distance between two opinions to induce assimilation or contrast are fixed in the SJ model and contextdependent in the meta-contrast model. This difference was not observable at the collective level in our simulations. We hypothesize this is mostly due to the opinion distribution used and the fact that the update process was iterated. For example, the fact that for SCT the whole context is important to evaluate (dis)similarity between individuals implies that two groups categorizing each other as outgroups may recategorize together in the presence of a third, very different group (see the end of paragraph 2.6). This behaviour is observable in the meta-contrast model after a single iteration, but not in the SJ model, unless a change of the latitude of acceptance or rejection is made between the two situations. However, repeated interaction under SJ model may lead to the same behaviour. In summary, further study would be needed to be able to put forward differences between the two models at the collective level, but our results suggest that the different types of convergence obtained by both models are the same, provided parameters are tuned accordingly. The MC model was developed to reproduce the categorization phenomenon and to identify prototypes associated with each category. As such, it reproduces correctly small-group interactions and may be more accurate than the SJ model in such situations. Results obtained at the macro level are also relevant and meaningful, but the SJ model seems to be more appropriate to describe large group, long term interactions, because it is much simpler.

4.3

Possible problems that could be studied using this model are the influence of the presence of explicitly formed groups and of particular spatial distribution of opinions (e.g. group of agents with initial extreme opinion, clustered or spread in the population). It should be possible to verify that the role of extremists is twofold: they can assimilate more central individuals and provoke large scale extremism in the population, or they can be rejected by others and reinforce the central consensus tendency among other agents or even "push" them toward an opposite extremism. Further, few empirical studies have attempted to measure behaviour at the individual level during large scale collective phenomena (e.g. demonstrations) due to the obvious difficulties to obtain results in such cases. However, many social psychological theories propose an explanation of these phenomena in terms of individual processes. This model could be used as a tool to obtain new experimental hypotheses related to such situations that would result from SCT.

Appendix: Prototypicality function in the case of a continuous distribution of opinions

A.1

In the beginning of this paper, we have presented the way the prototypicality function can be calculated given a discrete set X of initial opinions. We consider now a continuous (theoretical) initial distribution of opinions (x), which gives the density of opinion x for all $x \in [0;1]$. There is virtually no hypotheses to be made on this function; in particular it doesn't need to be normalized on [0;1] since this is automatically done in the calculus of d_{intra} and d_{inter} . In this case, the replacement of the sums by the corresponding integrals gives:

$$d_{\text{intra}}(x, \) = \frac{\int_{0}^{1} \rho(\xi) \mu(x,\xi) (x-\xi)^{2} \,\mathrm{d}\xi}{\int_{0}^{1} \rho(\xi) \mu(x,\xi) \,\mathrm{d}\xi}$$
(11)

and

$$d_{\text{inter}}(x, \) = \frac{\int_0^1 \rho(\xi) (1 - \mu(x,\xi)) (x - \xi)^2 \,\mathrm{d}\xi}{\int_0^1 \rho(\xi) (1 - \mu(x,\xi)) \,\mathrm{d}\xi} \ .$$
(12)

We then compute P as

$$P(x,) = a \cdot d_{inter}(x,) - (1 - a) \cdot d_{intra}(x,).$$
(13)

In the particular case of a uniform distribution ($(x) \equiv 1$), d_{intra} and d_{inter} are given by

$$d_{\text{intra}}(x) = \frac{w^2}{2} + \frac{w\left[\left(x-1\right)\exp\left(-\left(\frac{x-1}{w}\right)^2\right) - x\exp\left(-\left(\frac{x}{w}\right)^2\right)\right]}{\sqrt{\pi}\text{gerf}\left(\frac{x-1}{w}, \frac{x}{w}\right)}$$
(14)

and

$$d_{\text{inter}}(x) = \frac{1}{1 - \frac{w\sqrt{\pi}}{2} \text{gerf}\left(\frac{x-1}{w}, \frac{x}{w}\right)} \left\{ \left(x^2 - x + \frac{1}{3}\right) - \frac{w^2}{2} \left[\left(x - 1\right) \exp\left(-\left(\frac{x-1}{w}\right)^2\right) - x \exp\left(-\left(\frac{x}{w}\right)^2\right) \right] - \frac{w^3\sqrt{\pi}}{4} \operatorname{gerf}\left(\frac{x-1}{w}, \frac{x}{w}\right) \right\},$$
(15)

where $gerf(x_0, x_1) = \frac{2}{\sqrt{\pi}} \int_{x_0}^{x_1} exp(-\xi^2) d\xi$ is the generalized error function.

Acknowledgements

We would like to thank three anonymous referees for their helpful comments on a previous version of this paper.

Notes

1

We could model a discussion opposing two explicit groups (for example a political debate opposing two parties), by defining μ , for each group, as the characteristic function of this group.



AMBLARD, F. and Deffuant, G. (2004), The role of network topology on extremism propagation with the relative agreement opinion dynamics, *Physica A*, vol. 343, pp. 725–738.

BRAUER, M., Judd, C. (1996), Group polarization and repeated attitude expressions: A new take on an old topic, in W. Stroebe and M. Hewstone (Eds.), *European Review of Social Psychology*, vol. 7, John Wiley & Sons, chapter 6, pp. 173–207.

DAVID, B. and Turner, J. C. (1996), Studies in self-categorization and minority conversion: Is being a member of the out-group and advantage?, *British Journal of Social Psychology*, Vol. 35, pp. 179–199.

DAVID, B. and Turner, J. C. (1999), Studies in self-categorization and minority conversion: The in-group minority in intragroup and intergroup contexts, *British Journal of Social Psychology*, Vol. 38, pp. 115-134.

DEFFUANT, G., Amblard, F., Weisbuch, G. and Faure, T. (2002), How can extremism prevail? A study based on the relative agreement interaction model, *Journal of Artificial Societies and Social Simulation*, Vol. 5, No. 4, <<u>http://jasss.soc.surrey.ac.uk/5/4/1.html></u>.

DEFFUANT, G., Neau, D., Amblard, F. and Weisbuch, G. (2000), Mixing beliefs among interacting agents, *Advances in Complex Systems*, Vol. 3, pp. 87–98.

DITTMER, J. (2001), Consensus formation under bounded confidence, *Nonlinear Analysis*, Vol. 47, pp. 4615–4621.

HASLAM, S. A. and Turner, J. C. (1995), Context-dependent variation in social stereotyping 3 : Extremism as a self-categorical basis for polarized judgement, *European Journal of Social Psychology*, Vol. 25, pp. 341-371.

HEGSELMANN, R. and Krause, U. (2002), Opinion dynamics and bounded confidence models, analysis, and simulation, *Journal of Artificial Societies and Social Simulation*, Vol. 5, No. 3, .

ISENBERG, D. (1986), Group polarization: A critical review and meta-analysis, *Journal of Personality and Social Psychology*, Vol. 50, No. 6, pp. 1141-1151.

JAGER, W. and Amblard, F. (2005), Uniformity, bipolarisation and pluriformity captured as generic stylized behaviour with an agent-based simulation model of attitude change, *Computational and Mathematical Organization Theory*, Vol. 10, pp. 295-303.

KRAUSE, U. (2000), A discrete nonlinear and non-autonomous model of consensus formation, in S. Elaydi, G. Ladas, J. Popenda and J. Rakowski (Eds.), *Communications in Difference Equations*, Gordon and Breach Science Publishers, Amsterdam, pp. 227–236.

MACKIE, D. (1987), Systematic and nonsystematic processing of majority and minority persuasive communications, *Journal of Personality and Social Psychology*, Vol. 53, No. 1, pp. 41–52.

MCGARTY, C., Turner, J., Hogg, A., David, B. and Wetherell, M. (1992), Group polarization as conformity to the prototypical group member, *British Journal of Social Psychology*, Vol. 31, pp. 1–20.

OAKES, P. J., Haslam, S. A. and Turner, J. C. (1994), *Stereotyping and social reality*, Blackwell, Oxford, Cambridge.

OAKES, P. J., Turner, J. C. and Haslam, S. A. (1991), Perceiving people as group members : The role of fit in the salience of social categorizations, *British Journal of Social Psychology*, Vol. 30, pp. 125–144.

SALZARULO, L. (2004), Formalizing self-categorization theory to simulate the formation of

social groups, presented at the 2nd European Social Simulation Association Conference, Valladolid, Spain, 16th-19th September 2004.

TINDALE, R., Davis, J., Vollrath, D., Nagao, D. and Hinsz, V. (1990), Asymmetrical social influence in freely interacting groups: A test of three models, *Journal of Personality and Social Psychology*, Vol. 58, No. 3, pp. 438–449.

TURNER, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D. and Wetherell, M. S. (Eds.) (1987), *Rediscovering the social group : A self-categorization theory*, Blackwell, Oxford.

WATTS, D. and Strogatz, S. (1998), Collective dynamics of «small-world» networks, *Nature*, Vol. 393, pp. 440-442.

WEISBUCH, G., Deffuant, G. and Amblard, F. (2005), Persuasion dynamics, *Physica A*, Vol. 353, pp. 555-575.

Return to Contents of this issue

© Copyright Journal of Artificial Societies and Social Simulation, [2006]

