Reducing Prejudice: A Spatialized Game-Theoretic Model for the Contact Hypothesis

Patrick Grim[1], Evan Selinger[2], William Braynen[1], Robert Rosenberger[1], Randy Au[1], Nancy Louie[1] and John Connolly[1]

[1]Group for Logic and Formal Semantics, Dept. of Philosophy, SUNY at Stony Brook, Stony Brook NY 11794-3750
[2]Dept. of Philosophy, Rochester Institute of Technology, Rochester, NY 14623-5604
pgrim@notes.cc.sunysb.edu

Abstract

There are many social psychological theories regarding the nature of prejudice, but only one major theory of prejudice reduction: under the right circumstances, prejudice between groups will be reduced with increased contact. On the one hand, the contact hypothesis has a range of empirical support and has been a major force in social change. On the other hand, there are practical and ethical obstacles to any large-scale controlled test of the hypothesis in which relevant variables can be manipulated. Here we construct a spatialized model that tests the core hypothesis in a large array of game-theoretic agents. Robust results offer a new kind of support for the contact hypothesis: results in simulation do accord with a hypothesis of reduced prejudice with increased contact. The spatialized game-theoretic model also suggests a deeper explanation for at least some of the social phenomena at issue.

Introduction

There are a number of social psychological theories on the nature of prejudice (Adorno 1950, Campbell 1965, Tajfel and Turner 1986), but only one major theory of prejudice reduction: the contact hypothesis. According to the contact hypothesis, prejudice against members of one group by members of another will be reduced with increased social contact between members of the groups (Allport 1954). The hypothesis is simple and accords with common sense; it is understandable that it underlies a number of social policies, its most famous association being the desegregation of U.S. public schools (Patchen 1982, Schofield and Sagar 1977, Stephan 1978). Social psychological support for the contact hypothesis comes from laboratory, field, and survey methods (Cook 1985, Desforges et al. 1991, Robinson 1980, Sigelman and Welsh 1993, Stephan and Rosenfield 1978, Wilner et al. 1955). As with most large-scale social psychological hypotheses, however, there are practical and ethical obstacles to conducting large-scale controlled tests in which relevant variables can be manipulated. Those obstacles also impede the search for more fundamental explanation: if increased contact decreases prejudice, precisely how does it do so? As Pettigrew (1998) notes, the contact hypothesis itself does not address process. The attempts that have been made to understand mechanism, moreover, appeal to complex psychological processes of conceptual re-organization and the social dynamics of acquaintance and friendship (Brewer and Miller 1984, Gaertner et al. 1993, Pettigrew 1997).

We have found both a new type of confirmation for the contact hypothesis and hints toward deeper explanation in a game-theoretic simulation (Axelrod 1984, Epstein and Axtell 1996, Gilbert and Conte 1995, Schelling 1996).

A Minimal Model for Social Prejudice

Any model regarding prejudice in general must be capable of representing at least two different groups. In order to study prejudicial behaviors, as opposed to non-prejudicial, there has to be some range of behaviors that in some cases depend upon the group-identification of agent and recipient. Additionally, since prejudice has significant social effects, advantages and disadvantages can be expected to accrue depending on the behaviors that agents take and behaviors that are taken toward them. If prejudice is represented within the parameters of the contact hypothesis, moreover, changes in prejudicial behavior have to be analyzed with reference to circumstances of (a) contact and (b) lack of contact between members of at least two different groups.

These conditions dictate a minimal model using: (i) distinct groups, (ii) behaviors which may or may not be differentiated by actor and recipient groups, (iii) consequent advantages and disadvantages of those behaviors, (iv) some mechanism for updating patterns of behavior, and (v) conditions of greater and lesser contact between members of the groups. We think of the spatialized game-theoretic model used here as perhaps the simplest possible model of this form; little is built in beyond the minimal factors required for any model of prejudice adequate to the parameters of the contact hypothesis. An attempt at understanding ethnocentrism, with some points of contact with this model, appears as Axelrod 1997.

Agents are instantiated as cells in a 2-dimensional cellular automata array (Gilbert and Troitzsch 1999, Gutowitz 1990). Each cell interacts with only its 8 immediate neighbors—those cells touching it on sides and diagonals. Each cell is also of one of two colors—green or red—identifying its group. We can thus construct different conditions of contact by using arrays with different configurations of the two colors. Integrated contact can be modeled by randomizing the array by color, for example.

Segregation can be modeled by dividing the array into distinct color groups. This satisfies minimal conditions (i) and (v) above.

For the interaction between agents we have each cell play 200 rounds in an iterated Prisoner's Dilemma game with each of its 8 neighbors. We use the standard matrix. Each player gains 3 points for joint cooperation and 1 point for joint defection. Should one player defect and the other cooperate, the defector gets 5 points and the cooperator gets 0 (Table 1). The advantages and disadvantages of interaction in condition (iii) are reflected in each cell's total score. Here again we have constructed our model as simply as possible, using the standard *e. coli* game-theoretic model for conflict and cooperation, familiar from over 20 years of simulation research [3, 4].

		Player A	
		cooperate	defect
Player B	cooperate	3, 3	0, 5
	defect	5, 0	1, 1

Table 1. Standard Prisoner's Dilemma matrix, left gains to Player B.

We take as a basis just the 8 reactive strategies in an iterated Prisoner's Dilemma: strategies with behaviors on a given round determined only by behavior of the opponent on the previous round. These are shown in Table 2 using 0 for defect, 1 for cooperate, and a coding <i,c,d> to indicate a strategy's initial move *i*, its response *c* to cooperation by the opponent on the previous round , and its response *d* to defection by the opponent on the previous round.

<0,0,0>	All-Defect
<0,0,1>	Suspicious Perverse
<0,1,0>	Suspicious Tit for Tat
<0,1,1>	D-then-All-Cooperate
<1,0,0>	C-then-All-Defect
<1,0,1>	Perverse
<1,1,0>	Tit for Tat
<1,1,1>	All-Cooperate

Table 2. The 8 reactive strategies in an iterated Prisoner's Dilemma.

Suppose we start with a randomized cellular automata array of these 8 strategies. After 200 rounds of play with each of its neighbors, our cells total their scores. If a cell has a neighbor with a higher score, it adopts the strategy of its highest-scoring neighbor. In the case of a tie between higher-scoring neighbors, the strategy of one is chosen at random. This gives us a simple mechanism, well explored in the literature (Kennedy et al. 2001, Nowak and May 1993a, Nowak and May 1993b), as our updating schema for condition (iv). If we start with a randomized array of these 8 strategies, it is well known that dominance goes first to All-Defect and C-then-All-Defect, but that Tit for Tat (TFT) then grows in clusters and eventually conquers the entire array: a vindication for the robustness of TFT in a spatialized environment (Grim 1995, Grim 1996, Luna and Stefansson 2000).

Each of these 8 simple strategies is 'color-blind': each reacts to its opponent's previous play, but without regard to color. In order to meet condition (ii) in modeling prejudicial behavior, we add a single strategy PTFT ('Prejudicial Tit for Tat'). PTFT plays TFT with an opponent of its own color, but plays All-Defect against any opponent of the other color (Grim et al. 1998, Poundstone 1992).

In summary, we work with a 64 x 64 toroidal or 'wrap-around' array of 4096 cells, each of which has a background color of red or green. Different arrangements of those colors allow for different test conditions regarding the contact hypothesis. Each cell plays 200 rounds of an iterated Prisoner's Dilemma strategy with its 8 neighbors, following one of 8 'color-blind' strategies or a 'color-sensitive' strategy PTFT. After 200 rounds the gains and losses are totaled for each cell. If a cell has a higher-scoring neighbor, it adopts that strategy that has been most successful in its immediate neighborhood. Strategies are changed, but never colors, and strategy-updating is synchronous across the array. With a new configuration of strategies, we begin a new round of local play.

This is our minimal model for the conditions of the contact hypothesis. There are no complications of genetic algorithms or learning in neural nets, our agents do not construct any internal representations and indeed have no psychological depth at all. In Allport's original presentation, the contact hypothesis is qualified by a set of conditions that have been further elaborated and debated in the literature since: in order to reduce prejudice the contact at issue must be carried out by participants of equal status, who share common goals, participate in inter-group cooperation, and receive the support of authorities (1954). These complications are also largely missing in our model. While equal status for our cells is assured, cells operate in terms of purely individual gains and losses rather than common goals. Although there may be cooperation between individuals, there is nothing to model 'intergroup cooperation.' Since none of our cells represent authoritative figures, our model does not instantiate any kind of authoritative support.

Simulational Confirmation for the Contact Hypothesis

Despite the simplicity of the model, and despite the absence of the additional Allport conditions, our simulation robustly and persistently generates the phenomena predicted by the contact hypothesis. This suggests that the basic principles of contact networks and advantage, modeled in spatialialized game theory, may be sufficient to explain at least some aspects of the dynamics of real prejudice that have been noted in the social psychological literature.

First, consider an array that is carefully segregated in terms of background color. The array, divided in half down the middle, consists of green individuals on the one side and red individuals on the other (Figure 1).

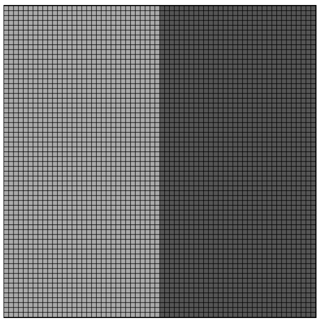


Figure 1. A segregated array of red and green

Over this array we layer a randomized distribution of strategies. A red cell might thus instantiate any of our 8 ‘color-blind’ reactive strategies <0,0,0>, <0,0,1>, ..., <1,1,1>, or might instead instantiate the color-sensitive strategy PTFT, representable as <1,1,0>/<0,0,0>. From that initial randomized array of strategies we follow the updating algorithm outlined: after playing 200 rounds of the iterated Prisoner’s Dilemma with each of its immediate neighbors, each cell surveys the success of its neighbors. If any has proven more successful, the cell copies the strategy of its most successful neighbor.

With a segregated background, the array converges within approximately 12 generations to a mixture of TFT and PTFT. The ‘prejudicial’ strategy, in other words, proves successful in occupying roughly 50% of the final array. In different runs, starting from different initial randomizations, either TFT or PTFT may show a slight dominance. The development of a typical array is shown in Figures 2 and 3.

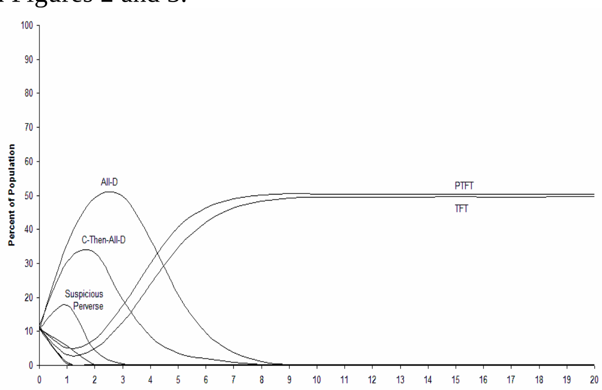


Figure 2. Percentages of the population for 9 strategies in an array segregated by color. 20 generations shown.

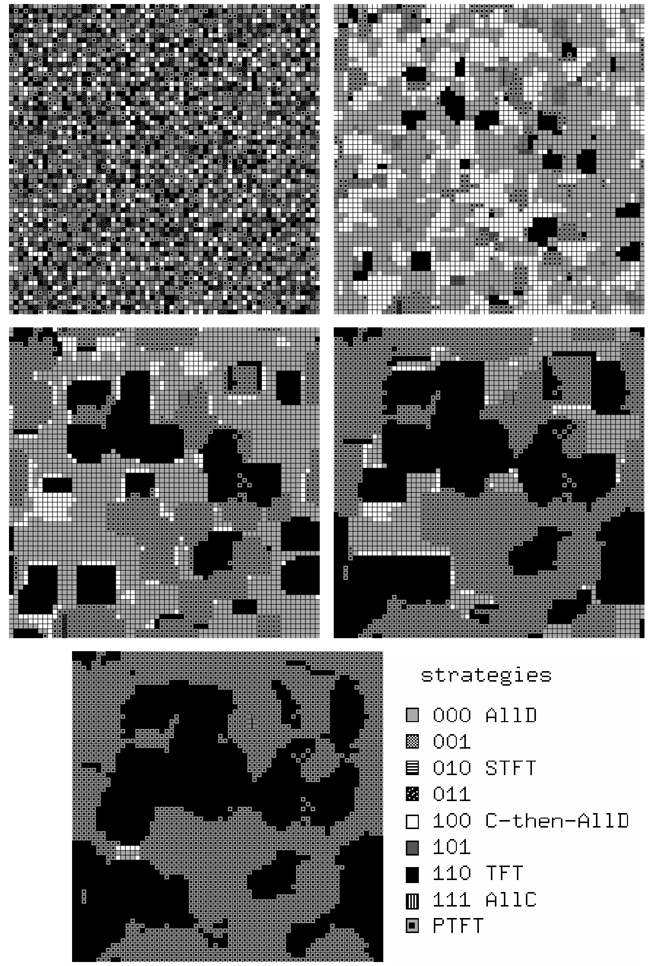


Figure 3. Evolution of randomized strategies to shared dominance by TFT and PTFT in an array segregated by color. Generations 0, 2, 4, 6, and 10 shown.

The claim of the contact hypothesis is that increased contact between groups will reduce prejudice. We therefore introduce a second array, with randomized background color (Figure 4).

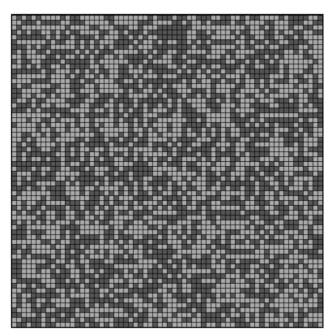


Figure 4. An integrated array of red and green

We overlay this integrated array with an initial randomization of our 9 strategies, as before, and repeat the simulation. Within 20 generations, the array shows a nearly complete conquest by TFT. Except for lone individuals or very small clusters, PTFT has been eliminated (Figures 5 and 6).

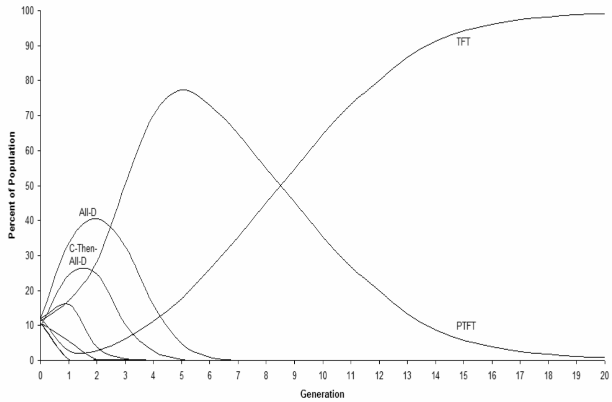


Figure 5. Percentages of the population for 9 strategies in an array randomized by color. 20 generations shown.

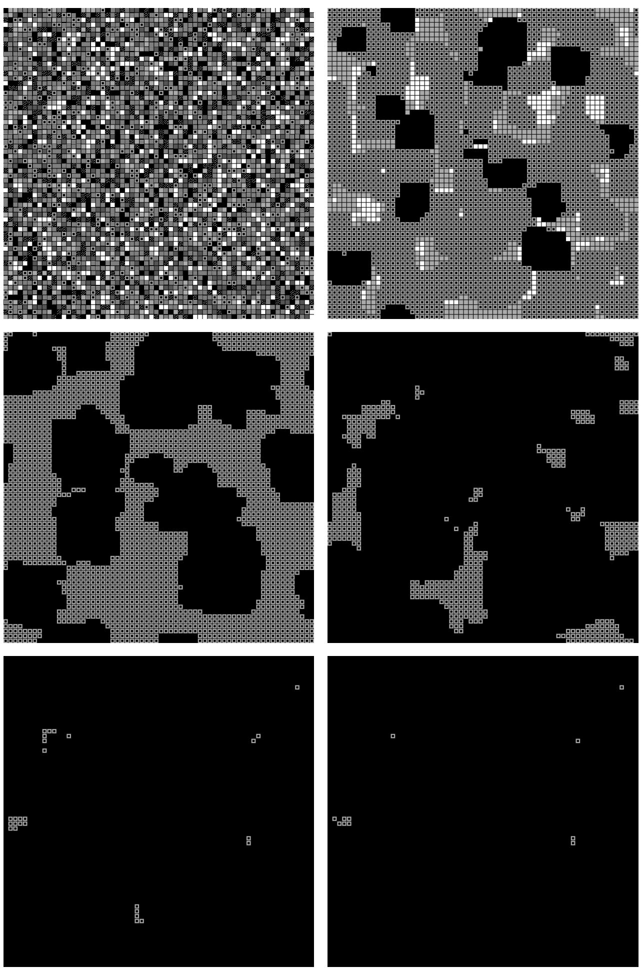


Figure 6. Evolution of randomized strategies to dominance by TFT in an integrated (randomized) color array. Generations 0, 4, 8, 12, 16, and 20 shown.

Social Identity Theory and a Stronger Result

We take this result to be a strong simulational instantiation of the basic phenomena predicted by the contact hypothesis. The result can be further strengthened by introducing a modeling factor borrowed from another theory of the nature of prejudice.

Social identity theory posits that much of one's identity is informed by the groups to which one belongs, and by the positive or negative perceptions of those groups. People are strongly motivated to develop a positive social identity; positive attitudes towards their own group and prejudice against others is one effect (Tajfel and Turner 1986). In our model PTFT is the only strategy that makes a distinction as to color. In order to model an additional value for 'social identification' we might then add a single point to PTFT cells when they are playing with neighbors that share the same color. A green PTFT playing a green All-C, for example, will be awarded 601 points instead of 600 points for 200 rounds of the iterated prisoner's dilemma.

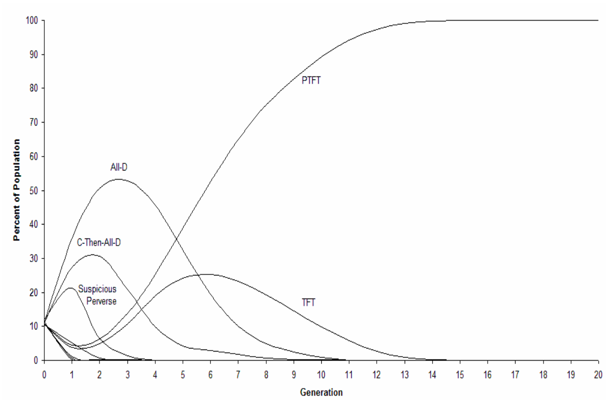


Figure 7. Percentages of the population for 9 strategies in a segregated array, with one extra 'social identification' point for PTFT playing a cell of its own color.

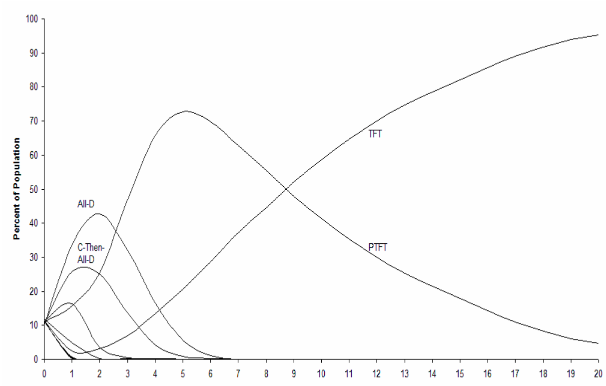


Figure 8. Percentages of the population for 9 strategies in an array randomized by color, with one extra 'social identification' point for PTFT playing a cell of its own color.

With this extra 'social identification' point for PTFT, the segregated array shown in Figure 2 now goes entirely to the prejudicial strategy PTFT (Figure 7). The array in which green and red strategies are mixed at random, on the other hand, still goes almost entirely to 'color-blind' TFT (Figure 8).

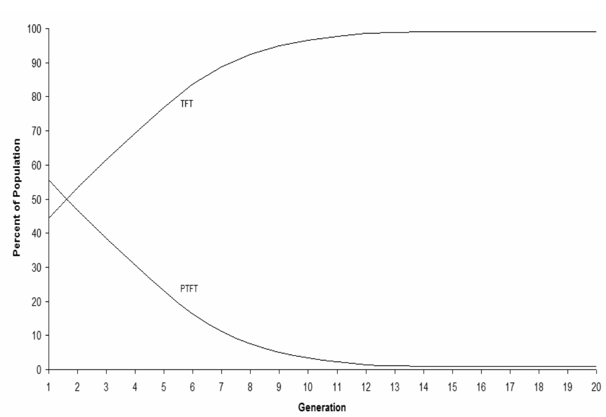
Eliminating established prejudice

Our models begin with a randomized array of strategies. In an environment mixed as to background color, the model shows evolution to dominance by TFT. In an environment segregated with regard to background color, the model shows evolution to co-dominance between TFT and PTFT in the simpler case, and evolution to full dominance by PTFT with a single additional 'social identification' point.

The contact hypothesis, however, is a hypothesis about prejudice reduction. What our models most directly show, it might be objected, is not reduction of established prejudice but the effect of contact in discouraging the spread of prejudice.

Many criticisms of model-building simply call for better models. In this case we can address the objection directly by starting not with a randomized array of strategies but with scattered territories of TFT and PTFT such as those shown in the final frame of Figure 2. What if we begin with this distribution of just these strategies, but with a mixed color background? Will established prejudice be eliminated?

The answer is 'yes'. Figure 9 shows evolution from such a distribution to clear dominance by TFT in a mixed environment. A similar shift to dominance by TFT can be shown if we start with an array dominated by PTFT except for very small patches of TFT, and give an additional 'social identification' point to PTFT when it plays its own color. Against a mixed color background PTFT is still progressively eliminated.



T

Figure 9. The Elimination of Established Prejudice: Triumph by TFT from scattered territories of PTFT and TFT in an array randomized by color.

Conclusion

Our attempt has been to construct a minimal model of prejudice adequate to the basic parameters of the contact hypothesis. The computationally interesting fact is that phenomena of precisely the sort the contact hypothesis would predict are evident in even this minimal model. In this respect, our results offer a model-confirmation of the contact hypothesis.

What our results further suggest is that patterns of individual advantage in different contact networks——captured by game-theoretical payoffs in a spatialized cellular automata—may be sufficient to explain why the contact hypothesis holds. Previous attempts to explain reduction of prejudice have appealed to complex psychological mechanisms of conceptual re-organization and the social dynamics of acquaintance and friendship (Brewer and Miller 1984; Gaertner et al.; Pettigrew 1997). Spatialized game-theoretic considerations of advantage and imitation seem to offer a simpler and deeper explanation for at least some of the phenomena at issue.

As aids to intuition and theoretical development, models such as this one may prove useful regarding other sociological and social psychological hypotheses that are difficult to test under strictly controlled conditions: hypotheses regarding deterrence and the death penalty, harm and pornography, or trickle-down economics.

It must be admitted that the methodological use of artificial societies is still at an early stage of development. Like both animal experimentation on the one hand and economic modeling on the other, simulational sociology has major limitations. Questions regarding the realism of a model and thus its generalizability to real societies inevitably remain. In this case, the matter is complicated by the fact that our model is so simple as to abstract away from many of the Allport provisos: the condition requiring equal status is satisfied, for example, but qualifications regarding common goals and intergroup cooperation play no role. Our results thus provide grounds for questioning whether these are in fact necessary for the basic phenomena of the contact hypothesis. Further research may be able to establish whether these conditions might play a more complicated role. Intergroup cooperation may not be strictly required for contact to reduce prejudice, for example, although certain types of intergroup conflict may work against the prejudice-reducing effect of contact.

References

Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., and Sanford, R. N. 1950. The authoritarian personality. New York: Harper.

Allport, G. W. 1954. The nature of prejudice. Cambridge, MA: Addison-Wesley.

Axelrod, R. 1984. The evolution of cooperation. New York: Harper.

Axelrod, R. 1997. The dissemination of culture: A model with local convergence and global polarization. Journal of Conflict Resolution 42(2): 203-226.

Brewer, M. B., Miller, N. 1984. Beyond the contact hypothesis: Theoretical perspectives on desegregation. In Miller, N., and Brewer, M. B. eds. The psychology of desegregation. New York: Academic Press, p. 281-302.

Campbell, D. T. 1965. Ethnocentric and other altruistic motives. In Levine, D. ed. Nebraska Symposium on Motivation. Lincoln, NE: Univ. of Nebraska Press, p. 283-311.

Cook, S. W. 1985. Experimenting on social issues: The case of school desegration. American Psychologist 40: 452-460.

Desforges, D. M., Lord, C. G., Ramsey, S. L., Manson, J. A., Van Leeuwen, M. D., West, S. C., & Lepper, M. P. 1991. Effects of structured cooperative contact on changing negative attitudes toward stigmatized social groups. Journal of Personality and Social Psychology 60: 531-44.

Epstein, J. M., & Axtell, R. 1996. Growing artificial societies—Social science from the bottom up. Cambridge, MA: MIT Press.

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachevan, B. A., Rust, M. C., 1993. The common group ingroup identity model: Recategorization and the reduction of intergroup bias. In Stroebe, W., & Hewstone, M. eds. European review of psychology 4, Chichester: John Wiley, p. 1-26.

Gilbert, N., and Conte, R. eds. 1995. Artificial societies: The computer simulation of social life. London: UCL Press.

Gilbert, N., and Troitzsch, K. G. 1999. Simulation for the social scientist. Philadelphia: Open Univ. Press.

Grim, P. 1995. Greater generosity in the spatialized Prisoner's Dilemma. Journal of Theoretical Biology 173: 353-359.

Grim, P. 1996. Spatialization and greater generosity in the stochastic Prisoner's Dilemma. BioSystems 37: 3-17.

Grim, P., Mar, G., & St. Denis, P. 1998. The philosophical computer. Cambridge, MA: MIT Press.

Gutowitz, H. ed. 1990. Cellular automata: Theory and experiment. New York: North-Holland.

Kennedy, J., Eberhart, R., and Shi, Y. 2001. Swarm intelligence. Menlo Park, CA: Morgan Kaufmann.

Luna, F., and Stefansson, B. 2000. Economic simulations in Swarm: Agent-based modeling and object oriented programming. Dordrecht: Kluwer.

Nowak, M. and May, R. 1993a. Evolutionary games and spatial chaos. Nature 359, 826-829.

Nowak, M. and May, R. 1993b. The spatial dimensions of evolution. Int. Journal of Bifurcation Chaos 3: 35-78.

Patchen, M. 1982. Black-white contact in schools: Its social and academic effects. West Lafayette, IN: Perdue Univ. Press.

Pettigrew, T. F. 1998. Intergroup contact theory. Annual Review of Psychology 49: 65-85.

Pettigrew, T. F. 1997. The affective component of prejudice: Empirical support for the new view. In Tuch, S. A. and Martin, J. K. eds. Racial attitudes in the 1990s: Continuity and Change. Westport, CT: Praeger, p. 76-90.

Poundstone, W. 1992. Prisoner's dilemma. New York: Doubleday.

Robinson, J. L. Jr. 1980. Physical distance and racial attitudes: a further examination of the contact hypothesis. Phylon 41: 325-332.

Schelling, T. C. 1996. Micromotives and macrobehavior. New York: Norton.

Schofield, J. W., and Sagar, H. A. 1977. Peer interaction patterns in an integrated middle school. Sociometry 40: 130-138.

Sigelman, L., Welsh, S. 1993. The contact hypothesis revisited: Black-white interaction and positive racial attitudes. Social Forces 71: 781-795.

Stephan, W. G. 1978. School desegregation: An evaluation of predictions made in Brown vs. Board of Education. Psychological Bulletin 85: 217-238.

Stephan, W. G. and Rosenfield, D. 1978. Effects of desegregation on racial attitudes. Journal of Personality and Social Psychology 36: 795-804.

Tajfel, H. and Turner, J. 1986. The social identity theory of intergroup behavior. In Wotchel, S., and Austin, W. G. eds. Psychology of intergroup relations. Chicago: Nelson, p. 7-24.

Wilner, D. M., Walkley, R., and Cook, S. W. 1955. Human relations in interracial housing: A study of the contact hypothesis. Minneapolis: Univ. of Minnesota Press.