

Modeling Prejudice Reduction:

Spatialized Game Theory and the Contact Hypothesis

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

1. Introduction

Philosophers have done significant work on concepts of ‘race’ and ‘racism’, on the ethics of a spectrum of race-conscious policies proposed to address a history of discrimination, and on identity and the experience of race (Outlaw 1996; Boxhill 2001; Bernasconi 2001a; Goldberg 1990; Appiah and Gutmann 1996). That work consists primarily of conceptual and normative analyses of prejudice and of the social policies designed to address it. Philosophers have also considered internal questions of racism within the canonical history of Western Philosophy, in for example Aristotle, Kant, and Hegel (Popkin 1980; Bernasconi 2001b).

What has been lacking, however, is sustained philosophical analysis regarding issues raised in the extensive social psychological literature: questions regarding the nature and formation of prejudice, questions regarding the social dynamics of prejudice, and questions regarding prospects for prejudice reduction. Here, explanation is central. If we cannot accurately explain how prejudice occurs and how it can be reduced, how are we to construct adequate public policy? The lack of philosophical attention in this area is thus particularly conspicuous and unfortunate.

As a first step toward remedying this situation—and with an eye toward public policy—we apply spatialized game theory and multi-agent computational modeling as philosophical tools: (1) for assessing the primary social psychological hypothesis regarding prejudice reduction, and (2) for pursuing a deeper understanding of the basic mechanisms of prejudice reduction. Social modeling in general has a philosophical pedigree that extends at least back to Hobbes and Locke. The particular techniques of social simulation employed here are relatively new, however, and raise important questions for the philosophy of science. For that reason we proceed reflexively, commenting throughout on both the promise of simulational techniques for social psychology and public policy and their inherent limitations.

2. The Social Psychology of Prejudice Reduction

There are a number of theories in the social psychological literature regarding the nature and sources of prejudice. Personality theories characterize prejudice in terms of personality types (Adorno et al. 1950; Eysenck 1954; Rokeach 1960), while dissociation theory portrays prejudice in terms of individual views that clash with what is culturally accepted (Devine 1989). Social identity theory and self-categorization theory depict prejudice as a natural result of the general human processes of categorization and group identification (Tajfel and Turner 1986; Turner and Oates 1989). The realistic conflict model views sees prejudice as a consequence of conflict between groups over limited resources (Campbell 1965), while relative deprivation theory emphasizes perceived resource differences (Gurr 1970).

Despite this range of theories regarding the nature of prejudice, there is only one 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). Allport qualifies the hypothesis with a set of conditions: 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. Details and elaborations of these qualifications continue to be debated (Pettigrew 1998), but the basic hypothesis is simple and accords with common sense. It is understandable that it underlies a number of social policies, its most famous association being *Brown v. Board of Education* fifty years ago and the desegregation of U.S. public schools (Schofield and Sagar 1977; Stephen 1978; Patchen 1982).

Support for the contact hypothesis comes from a range of studies. Cook is known for classic lab studies recreating contact behavior in the laboratory (1962, 1985). Other lab studies include Desforges' work with student groups and confederate outgroup members (Desforges, et. al. 1991) and Caspi's work on attitudes regarding the elderly (Caspi 1984). Fine offers confirmation for the contact hypothesis on the basis of archival records (1979).

Early survey studies in the field gauged attitudes of white seamen toward black seaman after the desegregation of the Merchant Marine (Brophy 1946), racial attitudes of residents in segregated and desegregated neighborhoods and housing projects (Deutsch and Collins 1951; Wilner et. al. 1955), and racial attitudes of Caucasian police officers working with African-American colleagues (Kephart 1957). More recent survey work has addressed close proximity and reduction in racial prejudice (Robinson 1985), the growth of positive racial attitudes (Sigelman and Welch 1993), attitudes of Anglo-Australians toward Cambodians, Hispanics, and Portuguese in small cities (Riordan 1987) and toward Vietnamese immigrants (McKay and Pittam 1993). Recent studies have also expanded the range of prejudice surveyed, including

prejudice against the old (Drew 1988), against computer programmers (McGinnis 1990), victims of AIDS (Werth and Lord 1992), the disabled (Anderson 1995), and homosexuals (Harek and Capitanio 1966).

As with most large-scale social psychological hypotheses, however, there are practical and ethical obstacles to conducting controlled tests in which the relevant variables can be manipulated. The ideal experiment would be one in which we start with large groups of people for whom we can establish a measure of prejudice against particular groups, and then move them to situations of (a) non-contact and (b) high contact with members of those groups. At a later date we would again measure prejudice in our population to establish the impact of contact as our independent variable. Of course that ideal experiment cannot be performed; the coercion required to stage such a study would be unethical even if feasible. A range of the studies above can nonetheless be seen as attempts at approximating such an ideal. In survey studies, data is used from existing contrast or comparison cases rather than from direct manipulation of a variable. Variables can be manipulated directly in laboratory studies, but only with small experimental groups and more limited contact.

Because existing studies do fall short of the ideal of direct variable manipulation in a large sample, questions of generalization will inevitably remain. Naturally-grown correlational data always carries the possibility of hidden and uncontrolled variables, and thus always leaves lingering doubts as to what causal pattern lies behind observed correlations. “In evaluating the contact hypothesis reviewers have noted that many studies have methodological inadequacies that make it difficult to determine what variables mediated any observed difference in attitude. The central problem is that it is generally unclear from the data whether the favorable attitude or contact came first” (Hewstone and Brown 1986, 12). Where there is confirmation for the

hypothesis in smaller groups, with artificial tasks, or in limited contact, doubt is bound to remain as to whether the effect will generalize to real contact situations involving large groups of people in everyday life. Where the available evidence addresses one type of prejudice regarding one type of group and in one type of context, it may also remain an open question whether the observed effects will generalize to all groups, to other contexts, and to other forms of prejudice (or indeed to individuals other than the subjects used in the study). “A persistent problem that dogs all such studies—whether dealing with integrated schools, housing, sports, combat conditions, work experience, or other settings—will the findings generalize?” (Hewstone and Brown 1986, xi).

The practical and ethical obstacles that block ideal testing of the contact hypothesis also impede the search for deeper explanation. If increased contact does decrease prejudice, precisely why does it do so? The contact hypothesis itself, as Pettigrew notes, does not address the question of mechanism (Pettigrew 1997, Zirkel and Cantor 2004). The attempts that have been made to understand mechanisms at issue, which are often addressed to qualifications of the hypothesis or to the question of how positive contact experiences generalize to attitudes about a larger group, have concentrated on high-level processes of cognitive organization and the social dynamics of acquaintance and friendship. Friendship with members of the other group is often cited as a process through which prejudice reduction occurs (Cook 1962; Pettigrew 1997). “Decategorization”, in which contact breaks down previous conceptual categories (Brewer and Miller 1984), could be understood as a mechanism of prejudice reduction, as could “recategorization”, in which existing categories are reorganized (Gaertner et. al. 1993). The consistent theme in such mechanisms is the idea that prejudice is reduced through individual learning; positive contact experiences somehow disconfirm previous conceptions of the

outgroup. Our research here suggests a distinctly different sort of mechanism that may also be at work.

In what follows, we provide an outline of a spatialized multi-agent game-theoretic model; this is an extremely simple artificial society, but one in which something like prejudice can be modeled and in which variables can be rigorously manipulated. The agents in our model are merely virtual agents, allowing us to avoid the practical and ethical obstacles of large-scale testing. In this regard, we follow modeling precedents in Schelling (1996), Epstein and Axtell (1996) and Axelrod (1997a, 1997b, 2003). Here, however, we apply these modeling techniques to issues of prejudice reduction and the contact hypothesis for the first time.

One of our claims is that data from modeling can offer a way to analyze the contact hypothesis. It cannot, of course, replace data from more traditional methods. Like both animal experimentation on the one hand and economic modeling on the other, simulational sociology of this sort has major limitations: questions regarding the realism of a model and thus its generalizability to real societies always remain. If viewed as a form of simulational experimentation, what is gained in large samples, repeatability, and manipulability of a wide range of variables must ultimately be paid for in work to show that the model is sufficiently realistic.

The model we offer is also intended to serve a further function. Simulations can offer a theoretical advantage by suggesting deeper lines of explanation: simple but deep mechanisms, for example, that may underlay perplexing surface complexity. This is one of the virtues claimed for artificial society work in both Schelling (1996) and Epstein and Axtell (1996). We think it is also a virtue of the model presented here. What the model suggests is that at least some aspects of contact and prejudice reduction can be understood at a level much simpler than

the cognitive and psychological mechanisms previously proposed in the literature. At least some of the phenomena captured by the contact hypothesis can be understood in terms of those factors of individual opportunity, advantage and disadvantage that are captured in simple spatialized game theory.

What we propose, then, is that the model outlined here has something to offer regarding both support for the contact hypothesis and deeper explanation. As an extremely simple model for prejudice, it offers an artificial laboratory suitable for further experimentation. It turns out that the phenomena predicted by the contact hypothesis appear robustly in this simple game-theoretical model. Virtual data, drawn from an environment of large numbers of individuals, and in which variables can be manipulated at will, can supplement our real data in important ways. Moreover, the fact that the phenomena predicted by the contact hypothesis appear even in a model this simple suggests that simple game-theoretic assumptions may be sufficient to explain at least some aspects of the phenomena at issue. What the model suggests is that there may be deeper explanations for some of the dynamics of prejudice and its reduction, operative across interactive groups but far simpler than the cognitive mechanisms usually appealed to. If so, of course, consideration of precisely these factors will be crucial to the directed and effective design of social policy.

3. The Modeling Background

Our model for prejudice is built on early work regarding spatialized game theory (Grim 1995; Grim 1996; Grim, Mar, and St. Denis 1998). We instantiate our agents throughout as cells in a 2-dimensional cellular automata array (Gutowitz 1990; Gilbert and Conte 1995; Gilbert and

Troitzsch 1999) (Figure 1). Interaction in the array is purely local: each cell interacts only with the 8 immediate neighbors touching it on each side and on the diagonals. The 64 x 64-cell array forms a torus, ‘wrapping around’ so that cells on the right edge have right-sided neighbors on the left edge, and cells at the top have upper neighbors on the bottom edge.

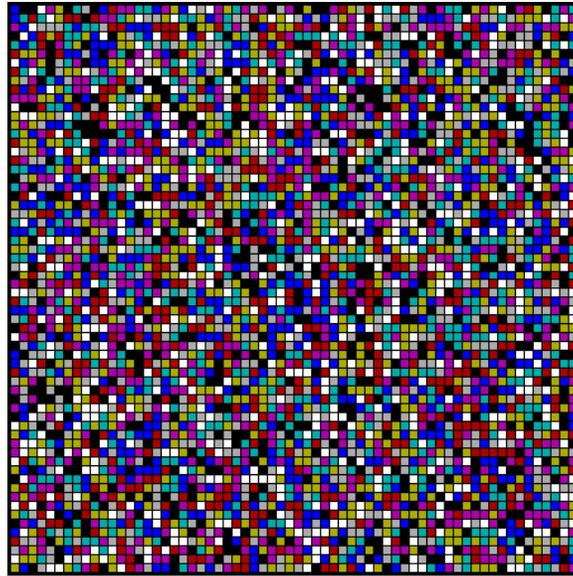


Figure 1 The 64 x 64 array.

The behaviors of our cells are defined in terms of the standard Prisoner’s Dilemma game. At each generation, we have our cells play 200 games with each of their 8 neighbors. We use the standard Prisoner’s Dilemma matrix; each player gains 3 points for joint cooperation and 1 point for defection. Should one player defect and the other cooperate, the cooperator gets 0 points while the defector gets 5 (Table 1).

		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 use as our basis for the prejudice model just the 8 reactive strategies in an iterated Prisoner's Dilemma: those strategies whose behavior on a given round is determined solely by the behavior of the opponent on the previous round. These basic strategies are shown as 3-tuples in Table 2 using 0 for defect and 1 for cooperate. Here a coding $\langle i, c, d \rangle$ indicates the strategy's initial move i (cooperate or defect), its move c if the opponent has cooperated on the previous round, and its move d if the opponent has defected on the previous round.

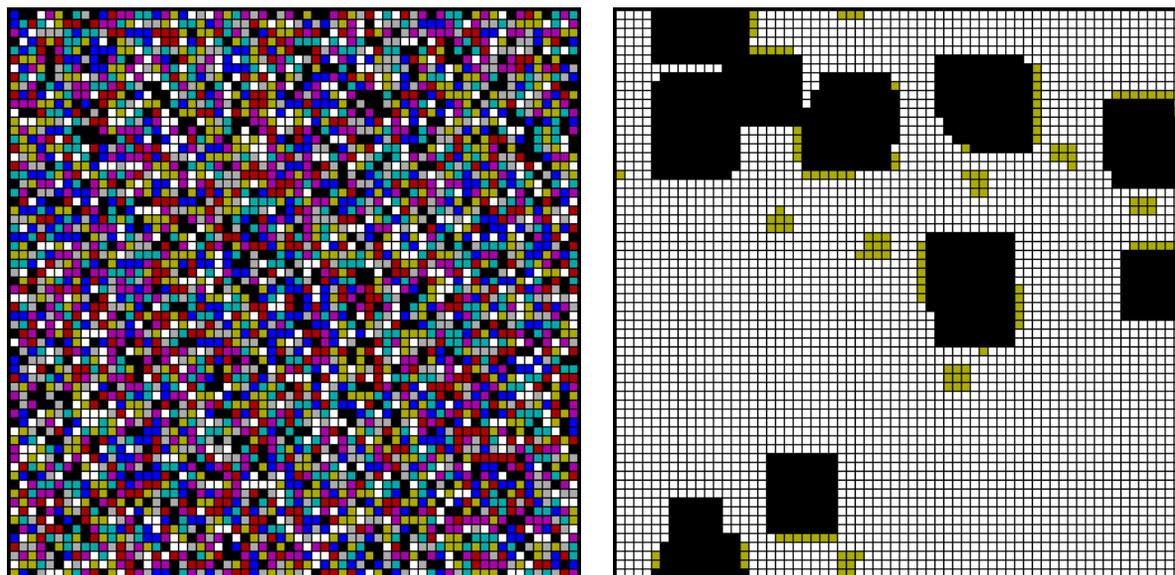
- $\langle 0,0,0 \rangle$ All-Defect
- $\langle 0,0,1 \rangle$ Suspicious Perverse
- $\langle 0,1,0 \rangle$ Suspicious Tit for Tat
- $\langle 0,1,1 \rangle$ D-then-All-Cooperate
- $\langle 1,0,0 \rangle$ C-then-All-Defect
- $\langle 1,0,1 \rangle$ Perverse
- $\langle 1,1,0 \rangle$ Tit for Tat
- $\langle 1,1,1 \rangle$ All-Cooperate

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

Suppose for a minute that we start with a randomized array of just these 8 reactive strategies.

After 200 rounds of iterated Prisoner's Dilemma games with its immediate neighbors, each cell totals its score. At that point, it looks around to see if any neighbor has garnered a higher total score; if so, it shifts its strategy to that of its most successful neighbor. Should no neighbor have a higher score, the cell retains its strategy. Should there be not just one higher-scoring neighbor but two tied with the highest score, the strategy of one is taken at random.

If we start with a spatialized array of just the 8 reactive strategies, it is well known that dominance goes first to a pair of exploitative strategies: All-Defect (All-D) and C-then-All-Defect (C-then-All-D). Once a range of vulnerable strategies have been eliminated, however, clusters of Tit for Tat (TFT) start to grow against the background of All-D and C-then-All-D. Eventually Tit for Tat conquers the entire array (Figure 2): a clear vindication in a spatialized environment of the general strengths of TFT (Grim, Mar, and St. Denis 1998).



(Figure 2)

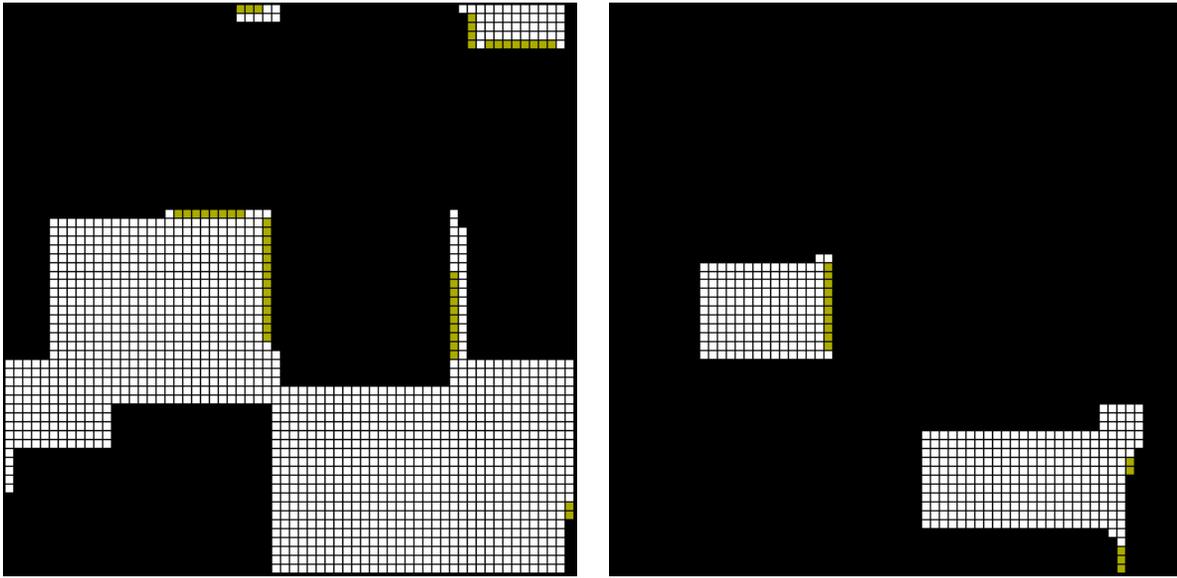


Figure 2 Conquest by TFT in a randomized environment of 8 reactive strategies.

For purposes of comparison, we will also consider a non-spatialized model throughout: the more familiar replicator dynamics. On this global model, we think of our population as divided into our 8 strategies, each with an (initially equal) percentage of the population. Play is global rather than local: on each generation, each strategy s plays 200 iterated games of Prisoner's Dilemma with each other strategy s' , multiplying its gain by the percentage of strategy s' in the population. A strategy's total for that generation is the sum of its weighted gains in play with all strategies in the pool. Strategy updating is global as well: more successful strategies increase their percentage in the population while less successful strategies lose percentage. At each generation the algorithm is repeated with revised percentages.

Symbolically, for a strategy s in the pool of strategies m , the proportion $P_{n+1}(s)$ of s on iteration $n+1$ is the ratio of $f_{n+1}(s)$ over the sum of $f_{n+1}(m)$ for all strategies m , where $f_{n+1}(s)$ for

each strategy s is the proportion of s at iteration n times the sum of values $V(s,m)$ to s of 200 games against each m weighted by the proportion of m at n .

$$P_{n+1}(s) = f_{n+1}(s) / \sum f_{n+1}(m) \text{ for all strategies } m$$

where $f_{n+1}(s)$ for any $s = P_n(s) \times \sum [(V(s,m) \times P_n(m))]$ for all strategies m .

A graph of the replicator dynamics for a population of just our 8 strategies again shows a triumph for TFT (Figure 3). In the case of the 8 reactive strategies, global replicator dynamics and the spatialized model give the same ultimate result. In more complicated models appropriate to prejudice and the contact hypothesis, however, it turns out that global and spatialized dynamics give radically different results.

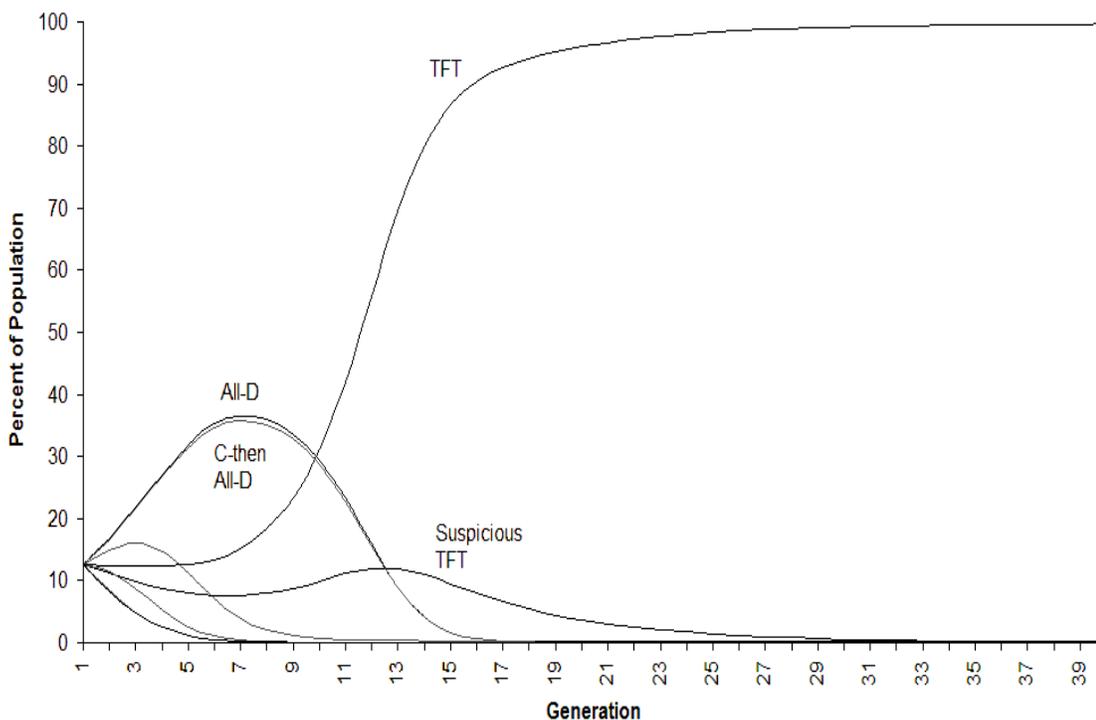


Figure 3 Global Replicator Dynamics (non-spatial) for 8 reactive strategies. 40 generations shown.

4. A Minimal Model for Prejudice

Our attempt is to construct a minimal model of prejudice adequate to the parameters of the contact hypothesis. Any model of prejudice must be capable of representing at least two different groups. In order to study prejudicial behavior as opposed to non-prejudicial behavior, we work with a range of possible behaviors. In some cases—the prejudicial cases—behavior is contingent on the group-identification of agent and recipient. In other cases—the non-prejudicial cases—behavior is ‘color-blind’ regarding groups. Because prejudice has significant social effects, we will want our agents to be advantaged or disadvantaged by at least some behaviors they take or that are taken towards them. If prejudice is to be represented within the basic parameters of the contact hypothesis, moreover, it has to be possible for changes in patterns of prejudicial behavior to occur in circumstances of contact, or lack of contact, between members of 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 or changing patterns of behavior, and (v) conditions of greater and lesser contact between members of the groups. We think of the spatialized game-theoretic model offered here as a very simple 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.¹

We instantiate our agents as cells in the 2-dimensional automata array outlined above, with interaction purely local: each cell plays only with its 8 immediate neighbors. In modeling

prejudice, however, each of our cells carries not only a strategy but a particular color—red or green—which define our social groups. We can therefore construct different conditions of contact between groups by configuring our array in different ways. A condition of segregated lack of contact can be produced by dividing the array down the middle, with green cells on one side and red cells on the other. A condition of mixed or integrated contact can be produced by choosing the color for each cell at random (Figure 4). Such a set-up satisfies conditions (i) and (v) for a minimal model, giving us two groups and different possible conditions of contact.

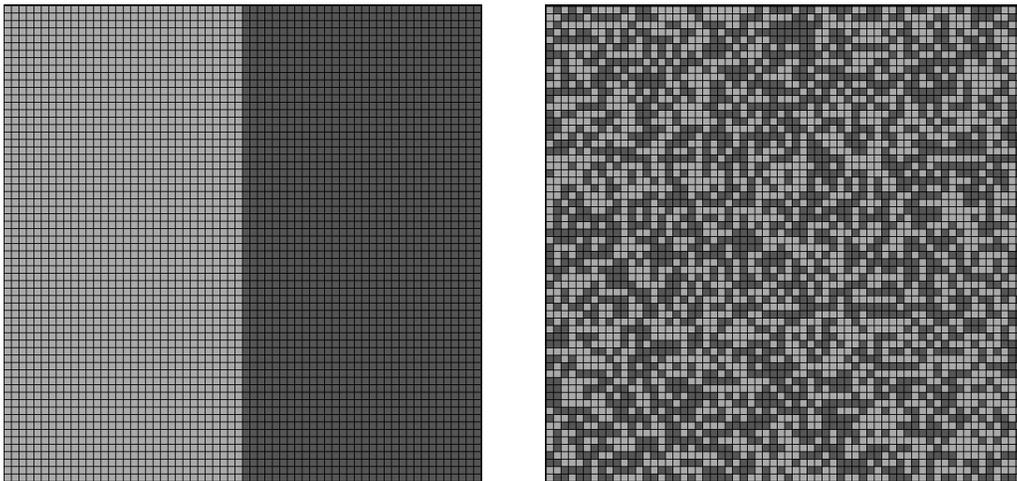


Figure 4 Segregated (left) and mixed patterns of background color.

Our behaviors are defined in terms of the standard Prisoner's Dilemma, using the standard matrix outlined above. The advantages and disadvantages to each agent required in condition (iii) are reflected in that cell's total score from interactions with its neighbors. Here as elsewhere, of course, the choice of gains and losses characteristic of the Prisoner's Dilemma is one of many alternatives (Skyrms 2001, 2004). Though we employ the game-theoretic model for conflict and cooperation that has been used in over 20 years of simulation work, we hope in

further work to explore the robustness of the effects at issue over changes in the game-theoretic matrix used.

Although the 8 reactive strategies constitute our modeling basis, we go beyond them in order to construct a model of prejudice. Each of the 8 simple strategies is ‘color-blind’: each reacts to its opponents’ previous play, but without regard to color. A green cell playing TFT, for example, plays the same way against a red cell as it does against another green cell. In order to meet condition (ii) in modeling prejudice, we add a single strategy PTFT (‘Prejudicial Tit for Tat’). PTFT plays TFT with any opponent of its own color, but plays All-Defect against any opponent of the other color. A green cell might thus instantiate any of our 8 ‘color-blind’ strategies $\langle 0,0,0 \rangle$, $\langle 0,0,1 \rangle$, ..., $\langle 1,1,1 \rangle$ or might instead instantiate the color-sensitive strategy PTFT, representable in self/other form as $\langle 1,1,0 \rangle / \langle 0,0,0 \rangle$.²

The updating mechanism required for condition (iv) is the same as in the simpler model. Each cell plays its strategy in 200 rounds with each of its neighbors, after which it totals its gains or losses. If a cell has a higher-scoring neighbor, it adopts that strategy that has proven most successful in its immediate neighborhood. It should be noted that strategies change, but never colors. With a new configuration of strategies, we begin a new generation of local play.

The model offered here can thus be characterized in terms of five basic conditions required for a minimal model adequate to prejudice and the parameters of the contact hypothesis. It incorporates very little else. The central question, of course, is whether the phenomena predicted for human societies in terms of the contact hypothesis—that prejudice will be reduced with increased contact—can be reproduced in a minimal model as simple as this one.

5. Support for the Contact Hypothesis

Our simulations show strong and robust results that parallel the contact hypothesis precisely.

We begin with an array that is carefully segregated in terms of background color, as in the left image in Figure 4. The array, divided in half down the middle, consists of green individuals on the one side and red individuals on the other. In addition to its color, each cell has a particular strategy. A red cell might thus also instantiate any of our 8 ‘color-blind’ strategies or might instead instantiate the color-sensitive strategy PTFT. The same is true of green cells. Strategies, as opposed to colors, we envisage as randomized across the cells of the array. If we look not at social group color but initial strategy distribution, our array has the appearance of Figure 1.

In each generation, each cell plays its neighbors in 200 rounds of the iterated Prisoner’s Dilemma. If its strategy is ‘color-blind’, its play does not depend on its own color or that of its neighbor. If its strategy is the color-prejudicial PTFT, on the other hand, it plays TFT with neighbors of its own color and All-D against players of the opposite color. At the end of a generation, cells copy the strategy of their most successful neighbor, as outlined above.

The evolution of a typical array in this segregated color environment, with an initially randomized distribution of strategies, is shown in Figure 5.

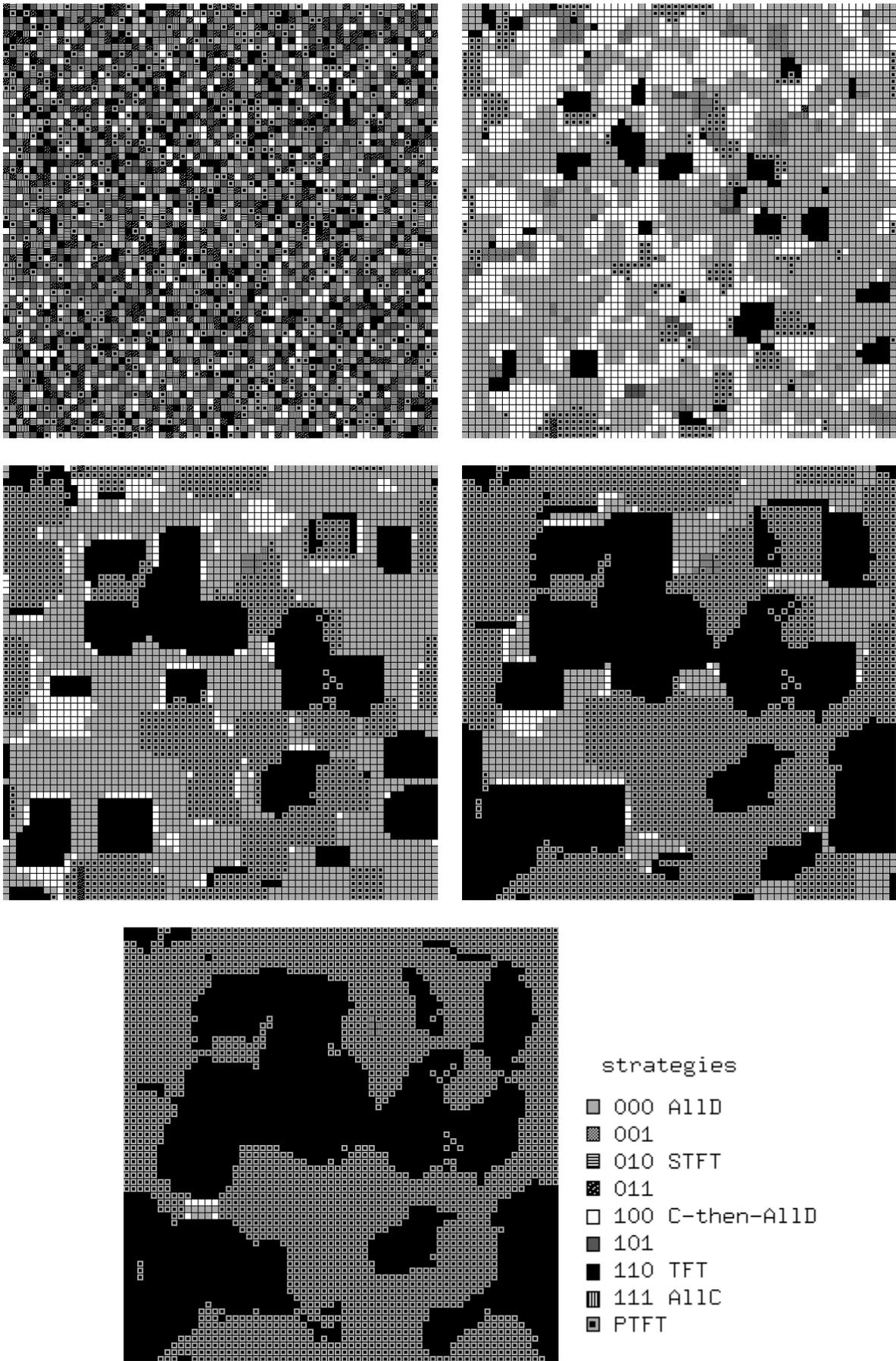


Figure 5 Evolution of randomized strategies to shared dominance by TFT and PTFT in an array segregated by color. Generations 0, 2, 4, 6, and 10 shown.

Within approximately 12 generations arrays typically converge to a mixture of TFT and PTFT, with no further changes except for random switches between TFT and PTFT by some cells along the color border.³ Prejudicial PTFT, in other words, proves successful in occupying roughly 50% of the final array. The same evolution graphed in terms of proportions of different strategies in the population as a whole is shown in Figure 6. There is, of course, a sensitivity to initial conditions in different initial randomizations; in some runs PTFT may occupy 60% of the array or so, in some runs only 40%.

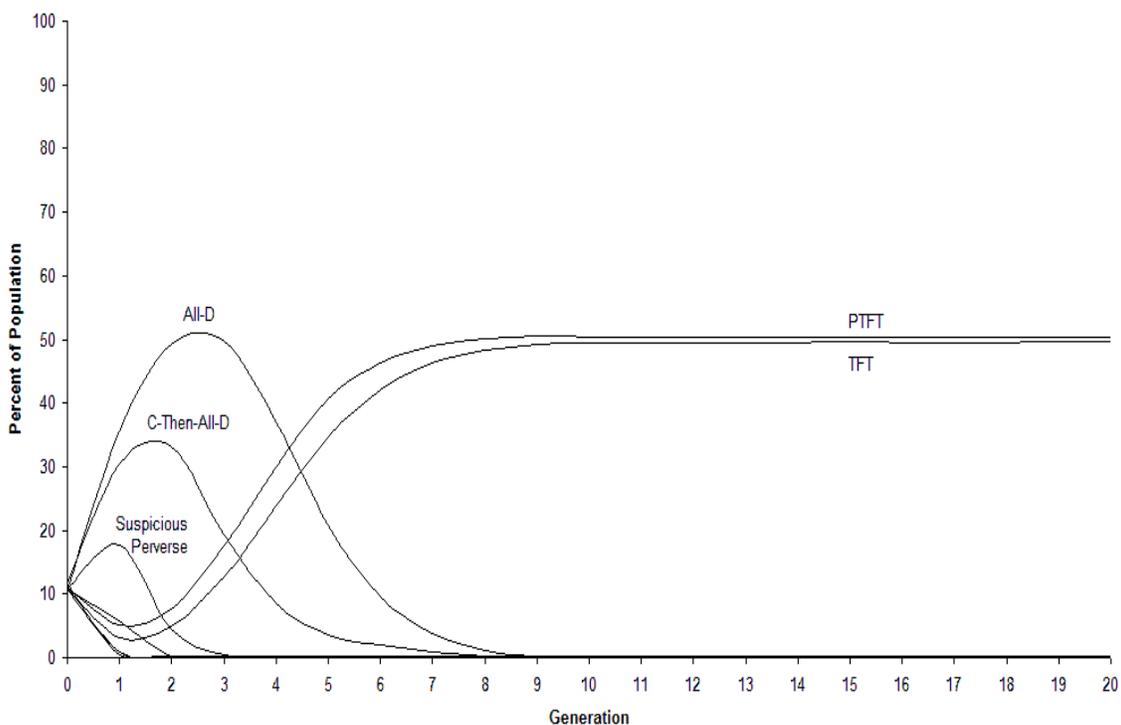
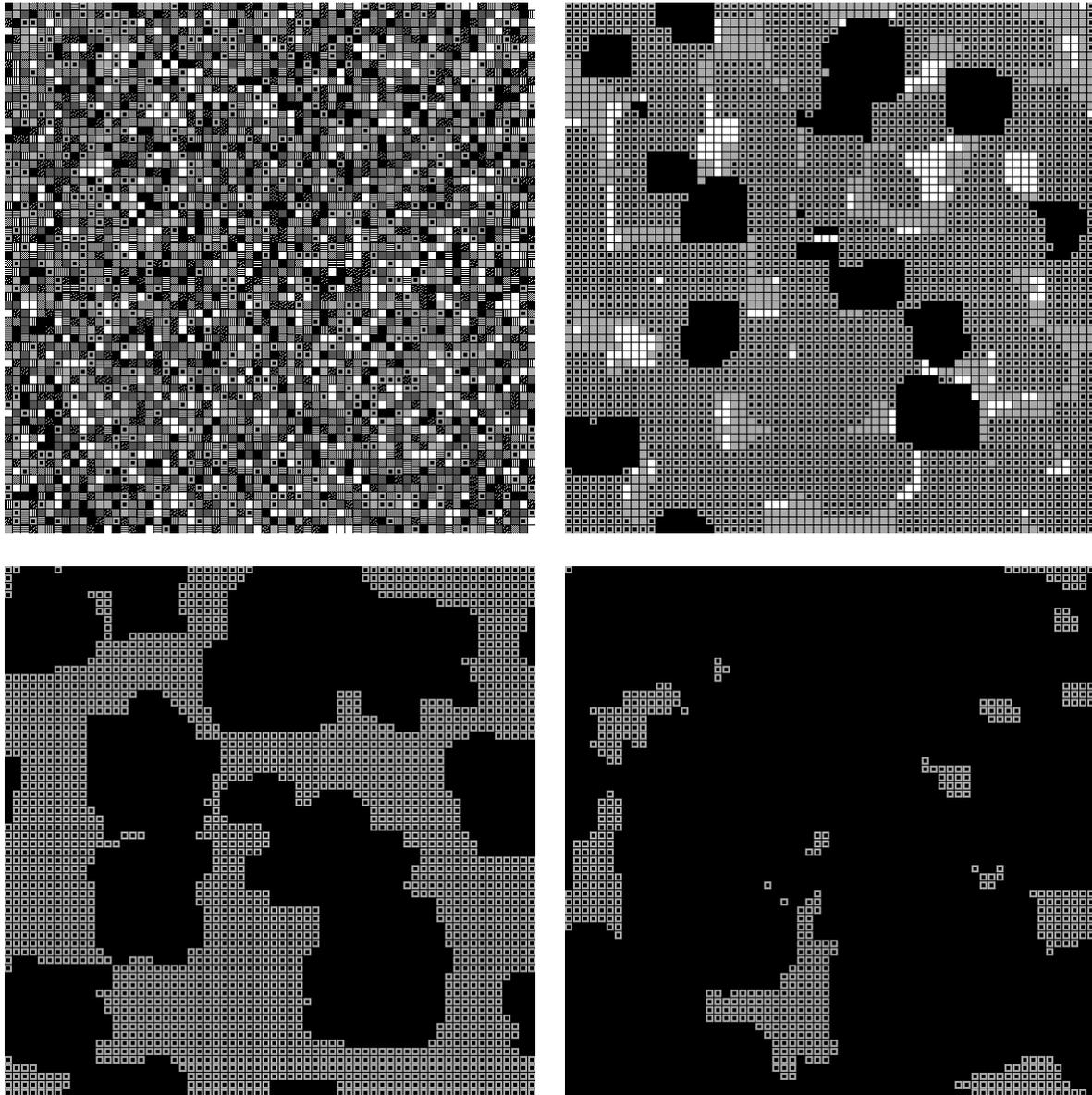


Figure 6 Percentages of the population for 9 strategies in an array segregated by color. 20 generations shown.

The claim of the contact hypothesis is that increased contact between groups will reduce prejudice. What then if we start with an array that is mixed rather than segregated with regard to

color? Here our background colors are randomized as in the right frame of Figure 4. Our strategies are randomized across the array as before. A typical evolution of the array gives us an importantly different result from before, shown in sequence in Figure 7 and in terms of proportions of the population in Figure 8.



(Figure 7)

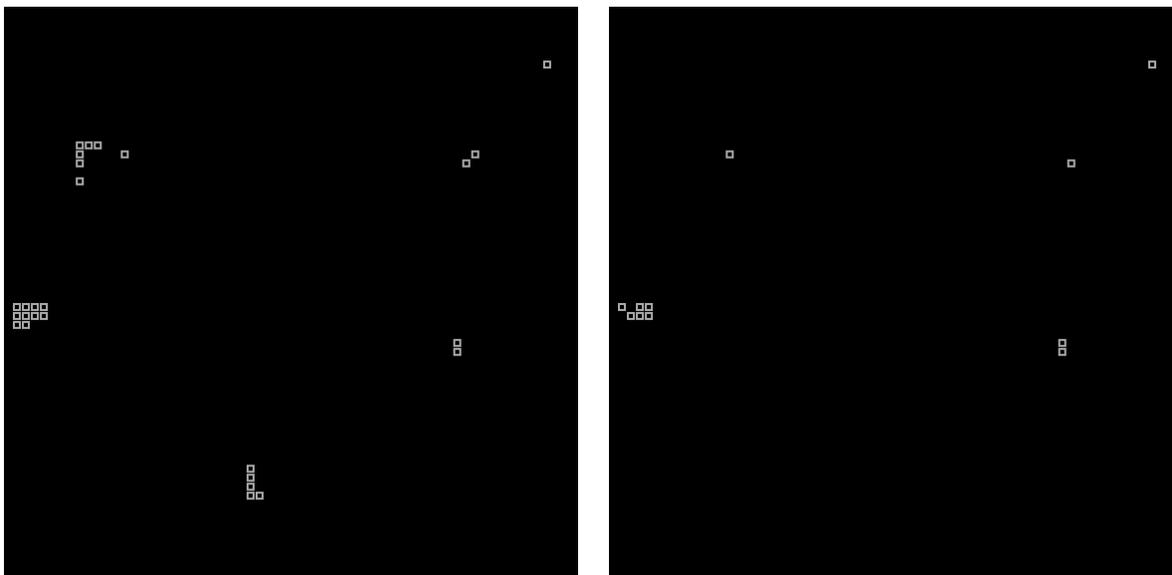


Figure 7 Evolution of randomized strategies to dominance by TFT in an integrated (randomized) color array. Generations 0, 4, 8, 12, 16, and 20 shown.

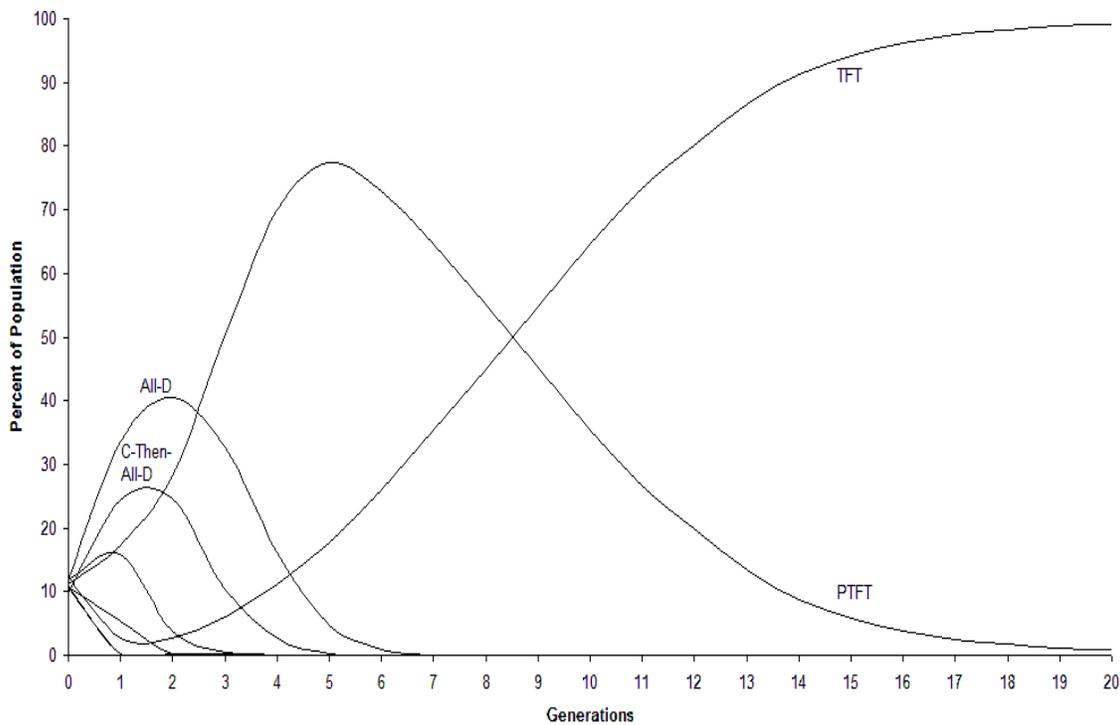


Figure 8 Percentages of the population for 9 strategies in an array randomized by color. 20 generations shown.

In a mixed array, TFT spreads to clear dominance, occupying all but very small clusters of PTFT in the array. When one examines where these PTFT clusters remains, one finds groups of cells of a single color clustered by chance in our color randomization of the array. Remnants of PTFT are thus the result of the same non-contact effect seen in the segregated case. In a pure checkerboard of colors, for instance, PTFT is eliminated completely.

We take this to be a strong simulational instantiation of the basic phenomena predicted by the contact hypothesis. A spatialized model using an iterated form of the standard Prisoner's Dilemma in local interaction, updating cells by imitation of successful neighbors, shows a spread of 'color-blind' TFT in mixed environments, but a 50-50 mix of TFT and the prejudicial PTFT in a segregated environment. Mixed environments in this simple simulation result in lower prejudice just as the contact hypothesis would predict for human societies. But here it is clear that simple game-theoretic dynamics are sufficient for the effect, without any appeal to complexities of human psychology.

It should be emphasized that the local action of our spatialized model, on the other hand, *is* crucial to this result. This can be shown by contrasting a global result using the standard replicator dynamics. Here too we can think of our population as divided between green and red individuals, but our 9 strategies are represented simply as (initially equal) percentages of the population. Each generation, play is global; each strategy plays against all strategies, weighting its gains against a particular strategy by that strategy's proportion in the population. As outlined in section 2, strategies with a higher total number of points then increase their percentage in the population, while strategies with lower totals decrease in the population; a strategy's current proportion is adjusted by multiplying it times the ratio of its average score over the average score in the population as a whole.

In this global model, we cannot of course compare a segregated ‘non-contact’ situation with a mixed ‘contact’ situation. Each individual plays globally against all strategies in the population, and all reproduction follows a global rather than a local algorithm. TFT is the clear winner using replicator dynamics (Figure 9). Use of a global model alone, therefore, would make the growth and spread of prejudicial strategies—shown in our segregated environment—incomprehensible. It is only with a model built in terms of the limited contact of spatial arrangement that the importance of prejudice appears. It is also in such a model that the phenomena of the contact hypothesis become evident.

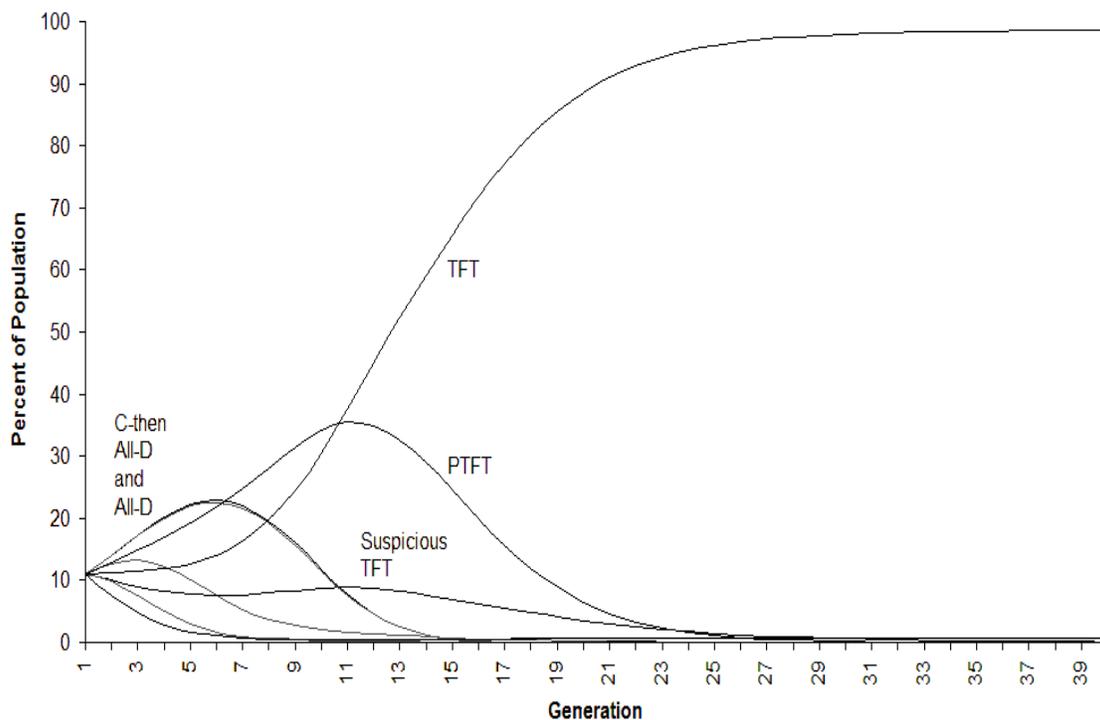


Figure 9 Global Replicator Dynamics (non-spatial): Conquest by TFT in a population of 8 strategies plus PTFT. 40 generations shown.

6. A Stronger Result

These results can be strengthened by introducing a modeling factor borrowed from a prominent 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, of course, 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 the total that PTFT cells gain in 200 rounds when they are playing with a neighbor of the same color. A green PTFT cell playing a green All-C, for example, would get 601 points instead of 600 as its total for 200 rounds.

With one extra 'social identification' point for PTFT, the segregated array now goes entirely to the prejudicial strategy PTFT (Figure 10). The array in which green and red strategies are mixed at random, on the other hand, still goes almost completely to 'color-blind' TFT (Figure 11).

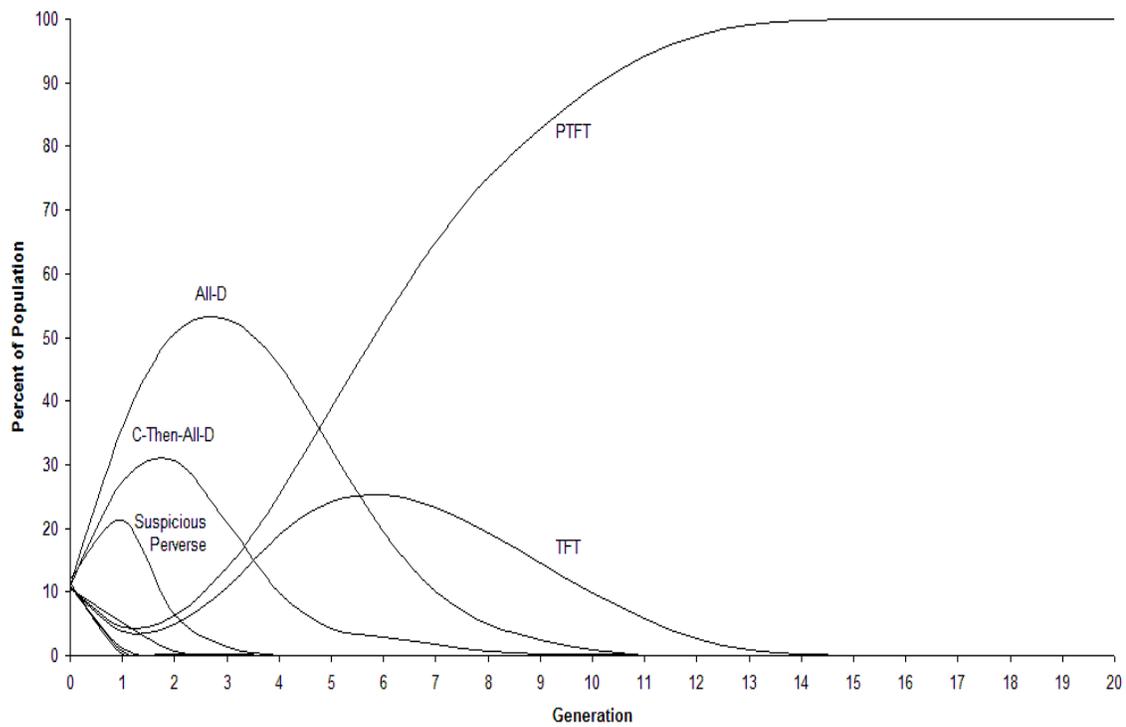


Figure 10 Percentages of the population for 9 strategies in an array segregated by color, with one extra ‘social identification’ point for PTFT playing a cell of its own color. 20 generations shown.

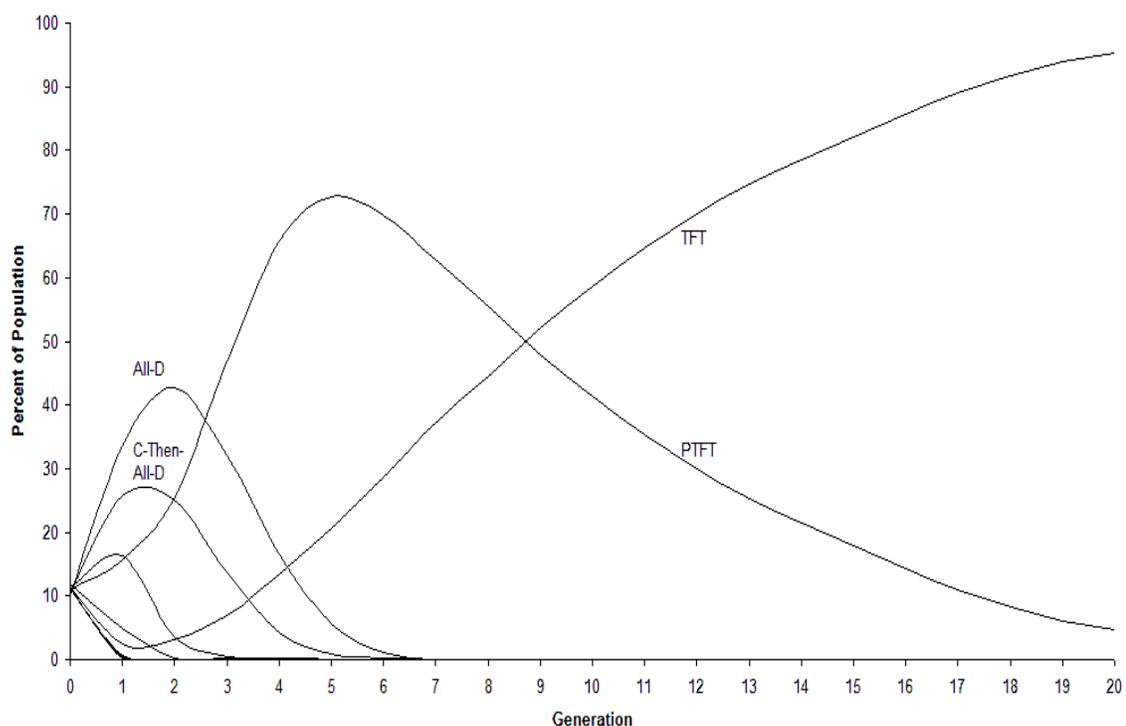


Figure 11 Percentages of the population for 9 strategies in an array randomized by color, with one extra ‘social identification’ point for PTFT playing a cell of its own color. 20 generations shown.

With one extra ‘social identification’ point, in other words, there is a full contrast in effects in a segregated and mixed environment. Color-blind TFT conquers in a mixed environment, prejudicial PTFT in a segregated environment.

Here again it should be emphasized that it is local action in a spatialized model that shows this clear contrast result. In a global model using replicator dynamics, on the other hand, with or without the ‘social identification’ point, conquest goes simply to TFT (Figure 12).

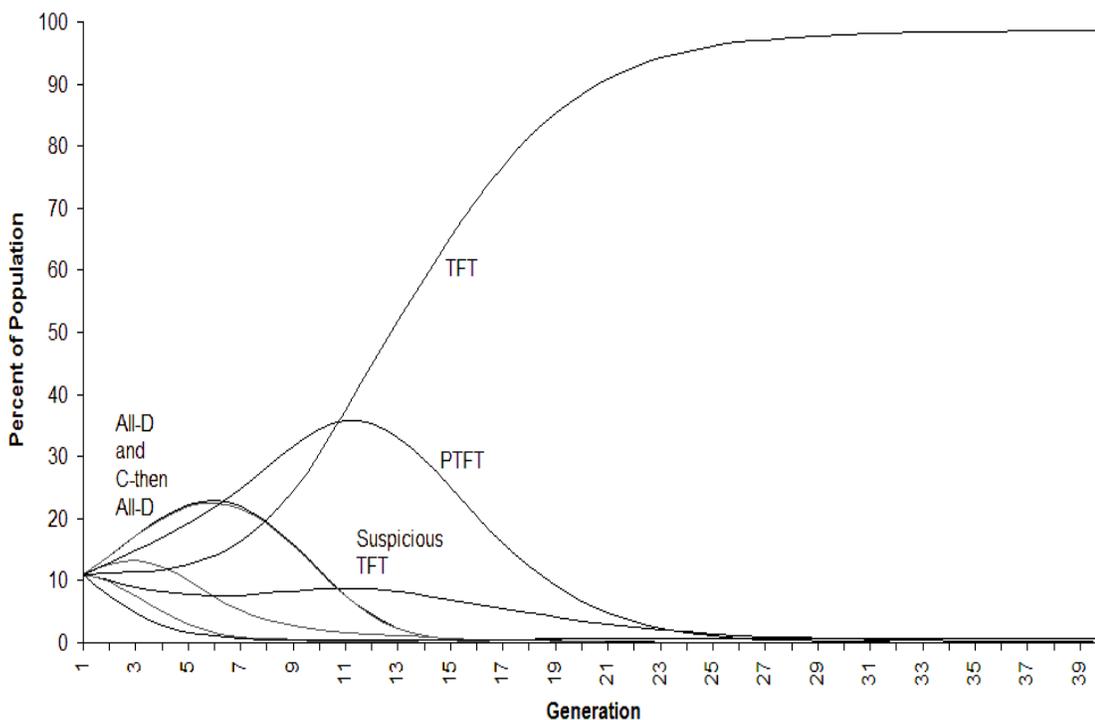


Figure 12 Global Replicator Dynamics (non-spatial): Conquest by TFT in a population of 8 strategies plus PTFT with an additional ‘social identification’ point. 40 generations shown.

7. Eliminating Established Prejudice

It might be objected at this stage that there is one respect in which our model differs from the ideal conditions of the contact hypothesis. What the results above show is evolution to joint dominance by TFT and the prejudicial PTFT from a randomized strategy pool in a segregated environment, contrasted with evolution to sole dominance by TFT in a mixed environment. But the topic of the contact hypothesis, strictly speaking, is reduction of prejudice once it is already there. It might be thought, therefore, that we should start with an initial situation not of a

randomization of all strategies but a distribution between prejudice and non-prejudice of the sort shown in the final array of Figure 5.

Such an objection can clearly be addressed by starting with a distribution purely of PTFT and TFT to which we have evolved in a segregated environment, and then changing the background colors to those of a mixed environment. This array again evolves to dominance by TFT (Figure 13). If we allot an extra point ‘social identification’ to PTFT when playing its own color, and moreover start with an array largely dominated by PTFT, the array still goes to clear dominance by TFT with a mixed color background (Figure 14).

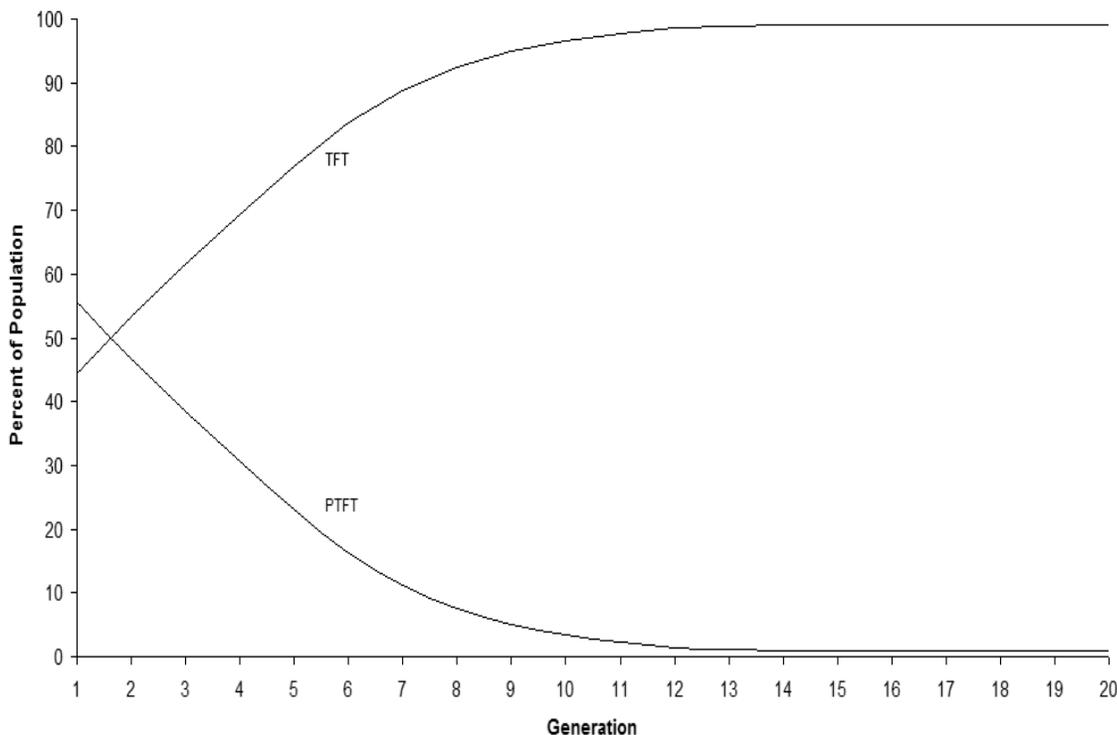


Figure 13. The Elimination of Established Prejudice: Triumph by TFT from scattered territories of PTFT and TFT in an array randomized by color.

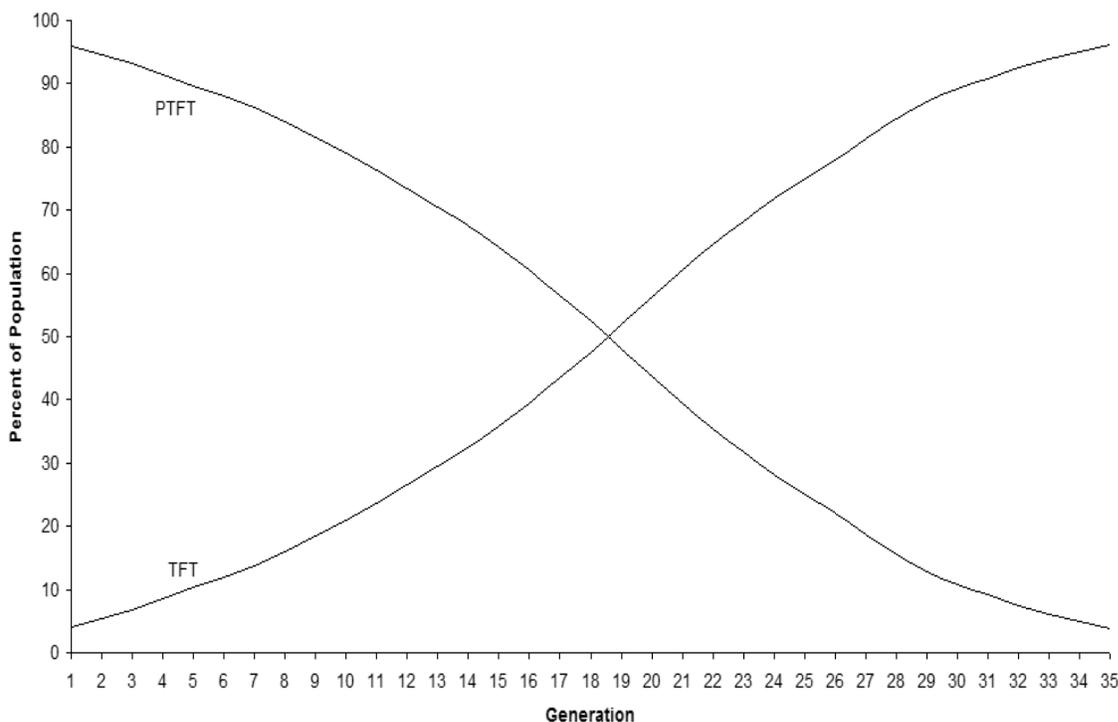


Figure 14. Triumph by TFT over territorial dominance by PTFT in an array randomized by color, even with one additional ‘social identification’ point for PTFT playing a cell of its own color. 35 generations shown.

In this model, then, contact both discourages the appearance of prejudice when starting from randomized strategies and reduces prejudice when it is already present.

Once our strategies have been simplified to TFT and PTFT, this effect—although striking—is not difficult to understand. Each cell could be playing with up to four types of neighbors: TFT of the same color, TFT of a different color, PTFT of the same color, or PTFT of a different color. In 200 rounds of the iterated Prisoner’s Dilemma, a TFT cell gains a total of 600 points as the reward for mutual cooperation with three of these four: TFT of either color and PTFT of its own color. Only in the case that it plays PTFT of a different color will TFT get a

score of 199, with 0 for its first play and 1 for each mutual defection from that point on. A PTFT cell, on the other hand, will get the rewards of mutual cooperation with only two of the four: PTFT of the same color and TFT of the same color. When playing a TFT of the other color, it will get 204 points—5 points for an initial defection against cooperation plus 199 points for mutual defection on the other rounds. When playing a PTFT of the other color, it will get a total of 200 mutual defection points.

In a segregated environment, one's neighbors are all of the same color. Gains for PTFT and TFT will be equal in that case, and we can expect them to occupy equal territory. In a mixed environment, on the other hand, and if the initial distribution of neighbor types are uniformly random, TFT has a clear advantage. TFT does well in play with 3 types of neighbors in a mixed environment, while PTFT does well in play with only 2. Given that difference in utilities, TFT can be expected to conquer.⁴

8. Complexities in an Expanded Model

In this section we complicate the basic model by expanding our pool of strategies. Here each of our strategies is a 'self-other' combination of our 8 original simple strategies. A cell with strategy $\langle 1,1,1 \rangle / \langle 0,0,0 \rangle$, for example, plays $\langle 1,1,1 \rangle$ (All-Cooperate) with cells of its own color, but plays $\langle 0,0,0 \rangle$ (All-Defect) with cells of the other color. With 8 possible strategies on the 'self' side, and 8 on the 'other', we have a population of 64 combinatory 'self-other' strategies. Of these, 8 are the original reactive strategies simpliciter: $\langle 1,1,0 \rangle / \langle 1,1,0 \rangle$ is pure color-blind TFT, playing the same strategy with cells of its own and the other color. $\langle 0,0,0 \rangle / \langle 0,0,0 \rangle$ is pure color-blind All-Defect. In place of a single color-sensitive strategy

PTFT, however, we would have a full 56 strategies that play differently depending on their own color and that of their opponent.

In this more complex environment, results become more complicated as well. Our results still support the thesis that a mixed environment strongly favors uniform play regardless of color, while a segregated environment allows strategies that distinguish between self- and other- to thrive. Promises of a deep and simple explanation for that phenomena remain. The complications that arise in this richer environment of strategies, however, also raise some questions about precisely what phenomena the model really instantiates and the extent to which that phenomena matches prejudice as we know it.

In an environment randomly mixed as to background color, evolution of a typical array is almost always to clear dominance by TFT, as shown in Figures 19 and 20.⁵ A few small clusters of $\langle 1,1,0 \rangle / \langle 0,0,0 \rangle$, $\langle 1,1,0 \rangle / \langle 1,0,0 \rangle$, $\langle 0,0,0 \rangle / \langle 1,1,0 \rangle$, and $\langle 1,0,0 \rangle / \langle 1,1,0 \rangle$ generally remain, contingent on the particular configuration of red and green cells. Where a randomization of green and red cells is replaced with a perfect checkerboard, evolution to dominance by TFT is total.

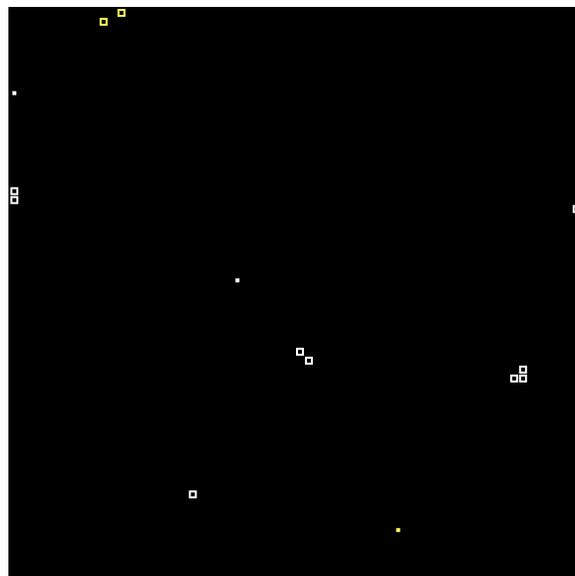


Figure 19 Dominance by TFT $\langle 1,1,0 \rangle / \langle 1,1,0 \rangle$ in mixed environment of 64 self-other strategies.

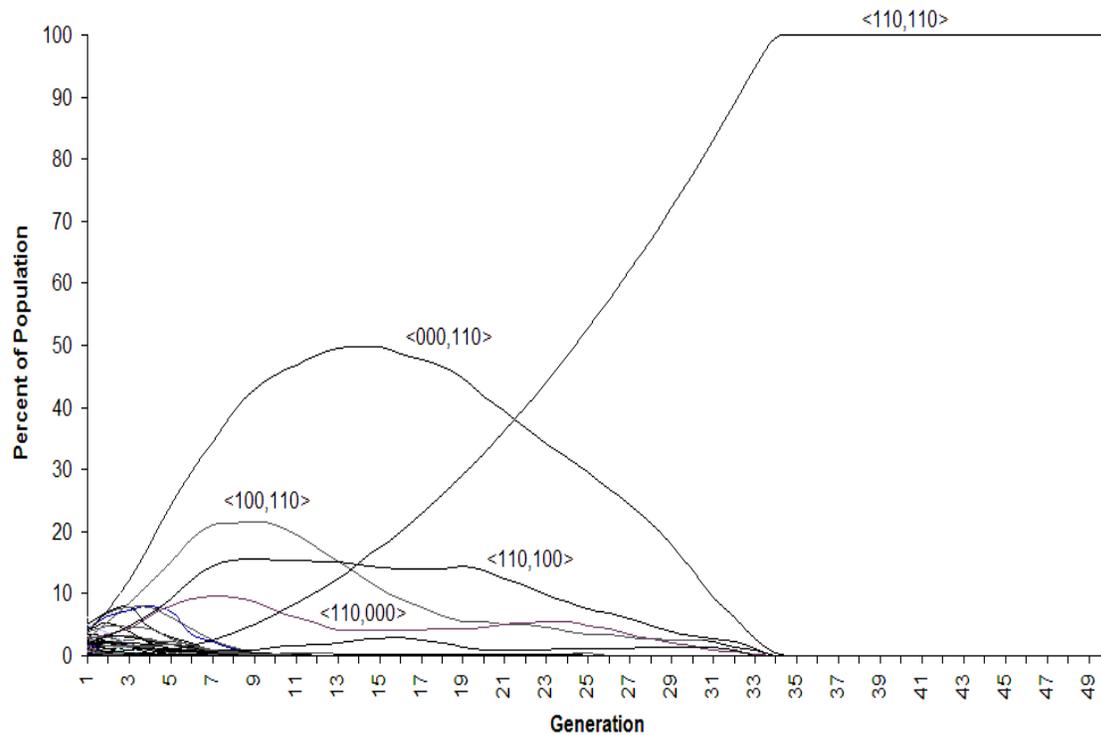


Figure 20 Evolution to dominance by TFT $\langle 110,110 \rangle$ in a mixed environment with 64 self-other strategies. 50 generations shown.

In a segregated array with 64 strategies, on the other hand, TFT does not evolve to dominance; territory is instead divided between 8 strategies. All of these play TFT, $\langle 1,1,0 \rangle$, with other cells of their own color. What strategy they play against cells of another color, however, varies across the field, from All-Defect $\langle 0,0,0 \rangle$ to All-Cooperate $\langle 1,1,1 \rangle$. Color-blind TFT occupies only about 1/8 of the field (Figure 21).

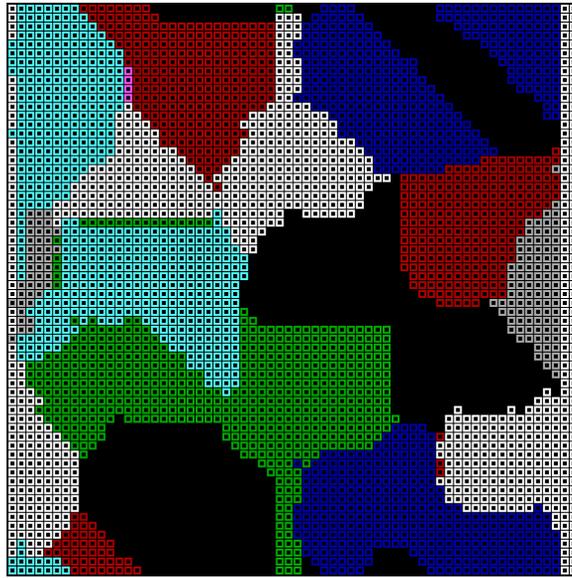


Figure 21 Shared dominance by 8 strategies with TFT core in segregated environment with 64 self-other strategies.

Here, as in the simpler studies, spatialization plays a major role. A version of this model that uses global replicator dynamics shows a simple dominance by TFT (Figure 22). In a global model, insensitive to local interaction, contrast effects between contact and non-contact environments thus simply become invisible.

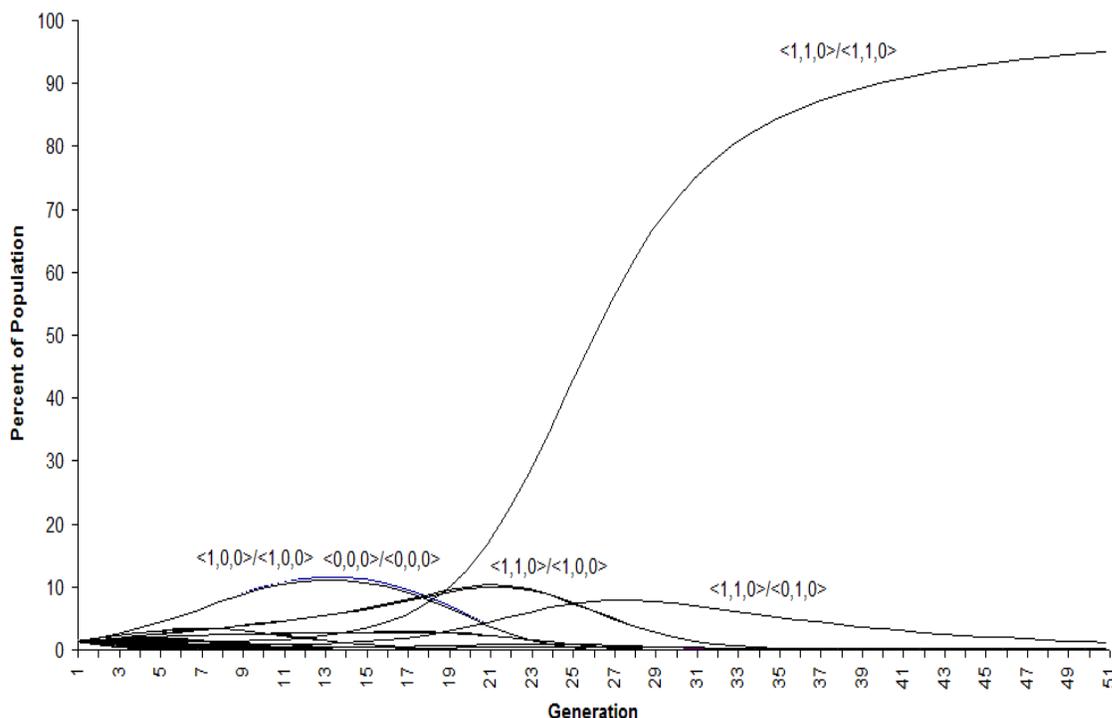


Figure 22 Global Replicator Dynamics (non-spatial): Conquest by TFT $\langle 1,1,0 \rangle / \langle 1,1,0 \rangle$ in a population of 64 strategies. 50 generations shown.

In this more complicated model, it remains true that color-blind TFT dominates in a mixed environment while color-discriminatory strategies flourish in a segregated environment. It is also clear in this more complicated model, however, that *any* strategy that plays TFT with its own color but which treats cells of the other color differently—whether it plays All-D against cells of the other color or All-C—flourishes equally with color-blind TFT in a segregated environment. The reason is clear: if a cell never interacts with others of another color, the strategy it specifies for play with them is irrelevant. Prejudice against another group (propensity to play $\langle 0,0,0 \rangle$, for example) and prejudice in favor of it (propensity to play $\langle 1,1,1 \rangle$) can flourish equally in an environment in which there is no contact with the other group at all.

Does this qualify our results, or the extent to which we can envisage them as supporting the contact hypothesis? As long as prejudice, by definition, involves an inherently negative view of other groups, the results do call for a qualification. For game-theoretic reasons alone, contact between groups can be expected to reduce differential treatment. That differential attitudes in realistic cases tend to be prejudicial *against* the other group is something that the game-theoretic mechanism alone does not explain. We might thus conclude that the mechanisms at issue can explain major aspects of the dynamics of prejudice, but not the full character of prejudice as we know it.

9. Conclusion

What we tried to construct is a minimal model of prejudice adequate to the parameters of the contact hypothesis. The results obtained within that model, we suggest, offer further support for the contact hypothesis. Perhaps more importantly, those results promise a deeper understanding of basic mechanisms that may be at play in prejudice reduction.

The results outlined above parallel precisely those that the contact hypothesis would predict for human groups: that increased contact results in the reduction of prejudice. In that sense, we regard our work as offering simulational support for the contact hypothesis. Simulational evidence of this kind, we want to suggest, can be a helpful supplement to data from more traditional methods involving real people and real societies. It certainly cannot replace real data, and questions regarding the applicability and generalizability of data from simulations will remain. Simulational data is in no sense unique in this regard; similar questions often remain regarding the applicability of economic models, for example, or the generalizability to humans of

medical experimentation on animals. Even the best of simulational evidence is far from conclusive. But very little of the data for social hypotheses can be said to be conclusive, from whatever source.

The best cases for confirming evidence for a general hypothesis drawn from simulational data will be ones in which the model can be shown to match independent data on limited cases. Epstein's work with general epidemiological models, calibrated to small but well-documented cases of the spread of smallpox, is an ideal to aspire to (Epstein and Axtell 1996). Simulations can support or undermine a hypothesis, however, even when that ideal isn't reached. Where the phenomena of a hypothesis prove impossible or difficult to produce in simulation, even when all the hypothesized factors of importance are modeled, we have reason to doubt the hypothesis. Consider our reaction to a Newtonian hypothesis of planetary motion, for example, were we unable to reproduce the result in simulation using all of the formulae at issue. Where the phenomena of a hypothesis prove 'fragile' in simulation—where it appears only for a very narrow window of relevant variables—we may also have grounds for suspecting that the hypothesis is unlikely to hold in any interesting variety of real environments. The fact that a phenomenon appears robustly and easily in simple modeled simulations, on the other hand, can be grounds for thinking that it may also appear in more complicated situations that include at least the basic variables of our model—the more complicated situations involving real people and real societies with which we are ultimately concerned. The phenomena offered here appear robustly and easily in precisely that sense.

Beyond a limited form of confirmation, however, simulations are noteworthy for the deeper explanations that they can offer. Because we have built them, and because we have built them simply, their structures are generally clear to us. If the results of a simple simulation can

then match some of the richness of the real effects that are their target, the model itself can be understood as suggesting an explanatory mechanism.

What is at issue here, of course, is not the explanatory power of models in general but how well game-theoretic resources can illuminate prejudice reduction in particular. If contact does reduce prejudice, precisely how does it do so? The contact hypothesis itself does not address the question of mechanism (Pettigrew 1998, Zirkel and Cantor 2004). Where attempts at understanding mechanism have been made in the social psychological literature, they have invoked high-level cognitive processing. Through acquaintance and friendship, ‘deategorization’ or ‘recategorization’, concepts regarding the outgroup are shifted. Roughly put, the core mechanism proposed is something like this: individuals learn from experience that members of the outgroup are not as bad as they had been stereotyped. What the simple simulations we have offered here suggest, however, is a radically different theory of the mechanism of prejudice reduction. Our agents operate at a level far below the cognitive sophistication required for friendship, categorization, stereotyping, or learning from experience. Our agents operate simply in terms of game-theoretic advantage and disadvantage, changing strategies to match those of their most successful neighbors. The fact that a model operating far below the cognitive level produces the phenomena of the contact hypothesis vividly and robustly suggests that there is a basic mechanism capable of reducing prejudice that is far simpler than has been supposed.

What our model suggests is that the simple elements of spatialized game theory—advantage, disadvantage, and imitation of the strategies of successful neighbors—may be sufficient to explain much of the effect of the contact hypothesis. In a segregated environment in which an agent deals with other agents of its own kind, there is no disadvantage to prejudicial

attitudes against agents of another kind. In an environment in which an agent deals with a mixed population of other agents, given some basic assumptions, prejudicial strategies prove disadvantageous. A major factor in the persistence of prejudice in segregated environments and its disappearance in conditions of contact may thus be the simple fact that patterns of advantage and disadvantage are different in those environments. A major reason why increasing contact between groups serves to reduce prejudice may simply be that it creates an environment in which prejudice becomes a disadvantageous strategy. What our model suggests, persistently and robustly, is that factors of advantage may go a long way toward explaining why contact can reduce prejudice.

We should make it clear that we are not claiming that all motivations are economic in nature or that all behavioral calculations are oriented toward advantage. It is in fact entirely consistent with our results to suppose that the cognitive mechanisms in agents that are more sophisticated than ours will involve more complicated attitudes. The fact that considerations of advantage are sufficient to produce the phenomena of the contact hypothesis in such a model, however, does seem to show that considerations of advantage can alone offer a significant explanation for major aspects of the phenomena in real cases. Also absent from our model is any emotional component to prejudice; it could thus be charged that the model leaves out something important about prejudice as we know it. Though it is clear that prejudice does standardly carry an emotional component, the question remains whether emotion is part of the primary mechanism of prejudice reduction. Emotion, though obvious in more complicated creatures, may be a secondary accompaniment to aspects of social dynamics that are captured in our model.

In Allport's original presentation, the contact hypothesis is qualified by a set of provisos that have been further elaborated and debated in the literature. Key provisos are that the contact

at issue must be between participants of equal status, sharing common goals, participating in groups that cooperate with one another, and that receive support from the authorities (Allport 1954; Brewer and Brown 1998; Pettigrew 1998). These complications are largely missing in our model. No differences in status are ever built into our model, and thus the condition of equal status is assured. It has also been held to be essential, however, that contact occurs under conditions of participation in cooperative tasks toward common goals (Brown 1988). Our cells operate purely in terms of individual gains and losses, and although there may be simulated cooperation between individuals, there is no larger cooperation between groups as wholes. Since our model lacks any hierarchy of authority, the ‘support of the authorities’ proviso is absent as well.

The social psychological research indicates that simple contact alone is inadequate to reduce prejudice. Our results do not in any way contradict that conclusion. In suggesting a deep mechanism for the phenomena of contact, however, our model does suggest new ways to understand the standard provisos. Here a bit of philosophical reflection is certainly not out of place. Among conditions relevant to any particular effect, we can certainly distinguish between (1) conditions that causally produce that effect, and (2) conditions that may inhibit or prevent the effect. Conditions causally productive for the motion of an automobile include the ignition of an air and gasoline mixture in its cylinders, and the transfer of power from the transmission to the wheels. Conditions that may block the motion of an automobile include solid walls in close proximity around it, or the force of a sudden and unexpected avalanche. ‘Defeating’ conditions of the latter kind can of course be multiplied at leisure: also adequate to block the motion of an automobile would be large busses in close proximity around it, large locomotives in close proximity around it, or large alien spaceships in close proximity around it. The category of

necessary conditions, however, indicating only conditions without which an effect will not appear, fails to distinguish between (1) causally effective conditions for an effect and (2) the absence of specific ‘defeaters’. Among the necessary conditions for an automobile’s motion are a properly running engine, a mechanical transfer of power to the wheels, and the absence of large buses, locomotives, and alien spaceships in close proximity around it. Only the first two, however, are causally productive of the motion.

In proposing pursuit of advantage in social exchanges as a driving force in prejudice reduction, our model suggests that the standard provisos associated with the contact hypothesis may play very different roles, despite the fact that all may qualify as necessary conditions. In at least some cases, those provisos may indicate conditions that facilitate or impede a central mechanism that we have identified in game-theoretic terms. For example, authorities might actively obstruct or undercut the prejudice reduction that would otherwise be expected from contact; negative intervention by authorities might function as a ‘defeater’. ‘Support of authorities’ or at least ‘lack of discouragement by authorities’ might then qualify as a necessary condition for prejudice reduction, but only in the sense that lack of alien spacecraft in close proximity qualifies as a necessary condition for automobile motion.

The model outlined here, we suggest, both offers simulational support for the contact hypothesis and promises a deeper understanding of important mechanisms of prejudice reduction. Because such a model can be manipulated at will, it also offers possibilities for further exploration. Here we use an identical number of red and green cells. How much can we change that condition and still have the contact effect appear? Will an integrated array with 60% red cells give us the same effect? With 70%? Equality in contact is perhaps foremost among the traditional provisos of the contact hypothesis, and equality is essentially built into the model we

have used here. All our cells carry the same ordered matrix with identical payoffs.

Heterogeneous variations on the model are also possible, however, and thus it should be possible to test the role of equality as a variable parameter. How equal do payoffs have to be in order for the contact effect to appear? Does the effect remain if we switch to a model in which the matrices of our individuals do not all qualify as Prisoner's Dilemma matrices, or not? A third promise of the kind of modeling we have outlined here is the possibility of exploring controlled variations on social hypotheses and proposals for social policy.

The general relevance of simulation to questions of public policy is clear. Axelrod's simulational work on the dissemination of culture (Axelrod 1997b), Axelrod and Hammond's models of ethnocentrism (Axelrod & Hammond 2003), Epstein's analyses of civil violence and genocide (Epstein 2002) and simulations regarding smallpox bioterrorism (Epstein, Cummings, Chakravarty, Singa, & Burke 2002) all offer prospects for understanding the dynamics of social phenomena and for testing public policy alternatives. What we have tried to offer here is a similar model, with similar promise for questions of public policy, regarding deep and crucial issues of prejudice.

References

- Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., and R. N. Sanford. 1950. *The authoritarian personality*. New York: Harper.
- Allport, G. W. 1954. *The nature of prejudice*. Cambridge, MA: Addison-Wesley.
- Anderson, L. S. 1995. *Outdoor adventure recreation and social integration: A social psychological perspective*. Ph.D. thesis. Univ. Minn., Minneapolis, MN.

- Appiah, K. A., and A. Gutmann. 1996. *Color Conscious: The Political Morality of Race*. Princeton Univ. Press.
- Anderson, L. S. 1995. *Outdoor adventure recreation and social integration: A social psychological perspective*. Ph.D. thesis. Univ. Minn., Minneapolis, MN.
- Axelrod, R. 1984. *The evolution of cooperation*. New York: Harper.
- Axelrod, R. 1997a. *The complexity of cooperation*. Princeton: Princeton Univ. Press.
- Axelrod, R. 1997b. "The dissemination of culture: A model with local convergence and global polarization," *Journal of Conflict Resolution* 42 (2): 203-226.
- Axelrod R., and R. A. Hammond. 2003. "The Evolution of Ethnocentric Behavior," Midwest Political Science Convention, April 3-6, Chicago, IL.
- Bernasconi, R. ed. 2001a. *Race*. Malden, MA: Blackwell.
- Bernasconi, R. 2001b. "Who Invented the concept of Race? Kant's Role in the Enlightenment Construction of Race." In *Race*, edited by R. Bernasconi, 11-36. Malden, MA: Blackwell.
- Brewer, M. B., and R. J. Brown. 1998. "Intergroup Relations," In *The Handbook of Social Psychology*, edited by Gilbert, D. T., Fiske, S. T., and G. Lindzey, vol. 2, 54-94. New York: McGraw Hill.
- Brewer, M. B., and N. Miller. 1984. "Beyond the contact hypothesis: Theoretical perspectives on desegregation," In *Groups in Contact: The psychology of desegregation*, edited by Miller, N. and M. B. Brewer, 281-302. New York: Academic Press.
- Brophy, I. N. 1946. "The luxury of anti-Negro prejudice," *Public Opinion Quarterly* 9:456-466.
- Brown, R. J. 1988. *Group processes: Dynamics within and between groups*. Oxford: Basil Blackwell.

- Campbell, D. T. 1965. "Ethnocentric and other altruistic motives," In *Nebraska Symposium on Motivation*, edited by Levine, D., 283-311. Lincoln, NE: Univ. of Nebraska Press.
- Caspi, A. 1984. "Contact hypothesis and interage attitudes: A field study of cross-age contact," *Social Psychology Quarterly* 47:74-80.
- Cook, S. W. 1962. "The systematic analysis of socially significant events: A strategy for social research," *Journal of Social Issues* 18:66-84.
- Cook, S. W. 1985. "Experimenting on social issues: The case of school desegregation," *American Psychologist* 40:452-460.
- Desforges, D. M., Lord, C. G., Ramsey, S. L., Manson, J. A., Van Leeuwen, M. D., West, S. C., and M. P. Lepper. 1991. "Effects of structured cooperative contact on changing negative attitudes toward stigmatized social groups," *Journal of Personality and Social Psychology* 60:531-44.
- Deutsch, M., and M. Collins. 1951. *Interracial housing: A psychological evaluation of a social experiment*. Minneapolis: Univ. Minn. Press.
- Devine, P. G. 1989. "Stereotypes and prejudice: Their automatic and controlled components," *Journal of Personality and Social Psychology* 56:5-18.
- Drew, B. 1988. *Intergenerational contact in the workplace: An anthropological study of relationships in the secondary labor market*. Ph.D. thesis. Rutgers Univ., New Brunswick, NJ.
- Epstein, J. M., 2002. "Modeling civil violence: An agent-based computational approach," *Proceedings of the National Academy of Sciences* 99: 7243-7250.
- Epstein, J. M., and R. Axtell. 1996. *Growing artificial societies—Social science from the bottom up*. Cambridge, MA: MIT Press.

Epstein, J. M., D. A. T. Cummings, S. Chavravarty, R. M. Singa, and D. S. Burke. 2002.

"Toward a Containment Strategy for Smallpox Bioterror: An Individual-Based Computational Approach," Working Paper No. 31, Center on Social and Economic Dynamics, The Brookings Institution, Washington D.C.

Eysenck, H. J. 1954. *The psychology of politics*. London: Routledge and Kegan Paul.

Fine, G. A. 1979. "The Pinkston settlement: An historical and social psychological investigation of the contact hypothesis," *Phylon* 40:229-242.

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

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

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

Goldberg, D. T., ed. 1990. *Anatomy of Racism*. Minneapolis: Univ. of Minnesota 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., and P. St. Denis. 1998. *The philosophical computer*. Cambridge, MA: MIT.

- Grim, P., Wardach, S., and V. Beltrani. 2003. "Location, Location, Location: The Importance of Spatialization in Modeling Cooperation and Communication." *Research Report #03-01, Group for Logic and Formal Semantics, Dept. of Philosophy, SUNY at Stony Brook.*
- Gurr, T. R. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.
- Gutowitz, H., ed. 1990. *Cellular automata: Theory and experiment*. New York: North-Holland.
- Herek, D. R., and J. P. Capitanio. 1996. "'Some of my best friends': Intergroup contact, concealable stigma, and heterosexuals' attitudes toward gay men and lesbians." *Personality and Social Psychology Bulletin* 22:412-424.
- Hewstone, M., and R. Brown. eds. 1986. *Contact and conflict in intergroup encounters*. Oxford: Blackwell.
- Kephart, W. M. 1957. *Racial factors and urban law enforcement*. Philadelphia: Univ. PA Press.
- Latane, B. 1996. "Dynamic social impact: The creation of culture by communication," *Journal of Communication* 46(4): 13-25.
- McGinnis, S. P. 1990. *Descriptive and evaluative components of stereotypes of computer programmers and their determinants*. Ph.D. thesis. New York: City Univ. NY.
- Mckay, S., and J. Pittman. 1993. "Determinants of Anglo-Australian stereotypes of the Vietnamese in Australia," *Australian Journal of Psychology* 45:17-23.
- Outlaw, L. T. 1996. *On Race and Philosophy*. New York: Routledge.
- Patchen, M. 1982. *Black-white contact in schools: Its social and academic effects*. West Lafayette, IN: Perdue Univ. Press.
- Pettigrew, T. F. 1997. The affective component of prejudice: Empirical support for the new view. In *Racial attitudes in the 1990s: Continuity and Change*, edited by Tuch, S. A. and Martin, J. K., 76-90. Westport CT: Praeger.

- Pettigrew, T. F. 1998. "Intergroup contact theory," *Annual Review of Psychology* 49:65-85.
- Popkin, R. 1980. "The Philosophical Bases for Modern Racism," In *The High Road to Pyrrhonism*, edited by R. Popkin, 79-102. San Diego: Austin Hill Press.
- Poundstone, W. 1992. *Prisoner's dilemma*. New York: Doubleday.
- Riordan, C. 1987. "Intergroup contact in small cities," *International Journal of Intercultural Relations* 2:161-185.
- Robinson, J. L. Jr. 1980. "Physical distance and racial attitudes: a further examination of the contact hypothesis," *Phylon* 41:325-332.
- Rokeach, M. 1960. *The open and closed mind*. New York: Basic Books.
- Rytina, S., and D. L. Morgan. 1982. "The Arithmetic of Social Relations: The Interplay of Category and Network" *American Journal of Sociology* 88:88-113.
- Schaller, M., and B. Latane. 1996. "Dynamic social impact and the evolution of social representations: A natural history of stereotypes," *Journal of Communication* 46(4): 64-77.
- Schelling, T. C. 1996. *Micromotives and macrobehavior*. New York: Norton.
- Schofield, J. W., and H. A. Sagar. 1977. "Peer interaction patterns in an integrated middle school," *Sociometry* 40:130-138.
- Sigelman, L., and S. Welsh. 1993. "The contact hypothesis revisited: Black-white interaction and positive racial attitudes," *Social Forces* 71: 781-795.
- Skyrms, B. 2004. *The Stag Hunt and the Evolution of Social Structure*. New York: Cambridge University Press.
- 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 D. Rosenfield. 1978. "Effects of desegregation on racial attitudes," *Journal of Personality and Social Psychology* 36:795-804.
- Tajfel, H., and J. Turner. 1986. "The social identity theory of intergroup behavior," In *Psychology of intergroup relations*, edited by Worchel, S. and W. G. Austin, W. G, 7-24. Chicago: Nelson.
- Turner, J. C. and P. J. Oates. 1989. "Self-categorization theory and social influence," In *The psychology of group influence*, edited by P. B. Paulus, 233-275. Hillsdale, NJ: Erlbaum.
- Werth, J. L. and C. G. Lord. 1992. "Previous conceptions of the typical group member and the contact hypothesis," *Basic Applied Social Psychology* 13:351-369.
- Wilner, D. M., Walkley, R., and S. W. Cook. 1955. *Human relations in interracial housing: A study of the contact hypothesis*. Minneapolis: Univ. Minnesota Press.
- Zirkel, S., and N. Cantor. 2004. "50 years after *Brown v. Board of Education*: The promise and challenge of multicultural education," *Journal of Social Issues* 60(1): 1-15.

Notes

¹ Latane and his colleagues have developed computer models in a research program regarding the movement of norms through culture (Latane 1996). Their research has many applications, including issues of prejudice (Schaller and Latane 1996). Latane's work provides nothing like a minimal model of prejudice reduction, however, and does not directly address the contact hypothesis.

² William Poundstone [1992] discusses a strategy in a non-spatialized context that he calls "Discriminatory Tit for Tat" (DTFT), which—like our PTFT—plays TFT with its own color

group but All-D with members of a different color group. Here Poundstone cites [Rytina and Morgan 1982], but we have been unable to find such a discussion in their work. In a spatialized context, [Grim, Mar, and St. Denis 1998] used ‘DTFT’ to designate a strategy which plays TFT with other DTFT cells and All-D with non-DTFT cells. In that context color was identified with strategy, and with that provision ‘DTFT’ fits Poundstone’s description. On further reflection, however, we think that the Grim-Mar-St. Denis ‘DTFT’, though interesting, is not what Poundstone intended. PTFT as outlined here is perhaps closer. We use ‘PTFT’ rather than ‘DTFT’ in order to avoid any further confusion.

³ For a cell with equally higher-scoring neighbors with strategies TFT and PTFT, which strategy it copies will be a matter of random choice. There are often cells in this predicament on the color border.

⁴ Simple calculations of relative utilities are not always enough, however, given the complications of a spatialized array. See Grim, Wardach, and Beltrani 2003.

⁵ This result does not hold in absolutely all cases, however. In order to flourish, TFT must have neighbors with the same or a complementary strategy; with 64 strategies in the pool, the probability of an initial critical mass of TFT is reduced. In the small number of cases in which TFT is not able to get an initial foothold the result is an array dominated in roughly equal portions by $\langle 1,1,0 \rangle / \langle 0,0,0 \rangle$, $\langle 1,1,0 \rangle / \langle 1,0,0 \rangle$, $\langle 0,0,0 \rangle / \langle 1,1,0 \rangle$, and $\langle 1,0,0 \rangle / \langle 1,1,0 \rangle$ alone.

In these mixed environment results it may seem peculiar that cells which play $\langle 0,0,0 \rangle / \langle 1,1,0 \rangle$ appear in roughly the same frequency as cells which play $\langle 1,1,0 \rangle / \langle 0,0,0 \rangle$. The latter play TFT with their own kind and All-Defect with others. But the former play All-Defect with their own kind, and TFT with the others. How can such a perverse strategy succeed? A

moment's reflection reveals that such a result is in fact to be expected. If a cell plays $\langle 1,1,0 \rangle / \langle 0,0,0 \rangle$ and has 4 neighbors of each color, it will be playing TFT with 4 and All-D with the other 4. If the cell had been playing a strategy inverted between 'self' and 'other'—strategy $\langle 0,0,0 \rangle / \langle 1,1,0 \rangle$ —it would still have been playing TFT with 4 of its neighbors and All-D with the other 4. A symmetry in the two cases is thus to be expected. If we vary the model so that each cell plays itself as well as its 4 neighbors, that symmetry disappears and the 'perverse' strategies disappear along with it.

The fact that both $\langle 1,0,0 \rangle$ and $\langle 0,0,0 \rangle$ variations appear in these configurations is purely a matter of the development of the array from initial random clusters. In a perfect checkerboard environment, $\langle 1,1,0 \rangle / \langle 0,0,0 \rangle$ invades and conquers $\langle 1,1,0 \rangle / \langle 1,0,0 \rangle$, for example.