

THE SOCIAL EVOLUTION OF GENOCIDE ACROSS TIME AND GEOGRAPHIC SPACE: PERSPECTIVES FROM EVOLUTIONARY GAME THEORY

CHARLES H. ANDERTON

Charles H. Anderton is W. Arthur Garrity, Sr. Professor in Human Nature, Ethics and Society and Professor of Economics, Department of Economics and Accounting, College of the Holy Cross, Worcester, MA, USA. A former co-editor of *Defence and Peace Economics*, he may be reached at canderto@holycross.edu.

Abstract

A standard evolutionary game theory model is used to reveal the interpersonal and geographic characteristics of a population that make it vulnerable to accepting the genocidal aims of political leaders. Under conditions identified in the space-less version of the model, genocide architects can engineer the social metamorphosis of a peaceful people-group into one that supports, or does not resist, the architects' atrocity goals. The model reveals policy interventions that prevent the social evolution of genocide among the population. The model is then extended into geographic space by analyzing interactions among peaceful and aggressive phenotypes in a Moore neighborhood. Key concepts of the analyses are applied to the onset and spread of genocide during the Holocaust (1933-1945) and to the prevention of genocide in Côte d'Ivoire (2011).

Breeders artificially select animal qualities by manipulating the animals' DNA, interactions with other animals (e.g., mating, artificial insemination), and environment (e.g., diet, healthcare). In so doing, the breeders create animal phenotypes (traits) that they deem desirable. The evolution of the animals is relatively quick, brought about as it is, not by natural selection, but by artificial selection. Metaphorically, the architects of genocide and other forms of mass atrocity seek to create (or in a sense "breed") certain types of social outcomes by refashioning how people interact with each other in society and by artificially manipulating peoples' perceptions of themselves, others (including the genocidal architects and people targeted for extermination), and the environments (institutions, perceptions of history, cultural norms, etc.) in which they live. The social outcomes the genocide architects seek to engineer are those that advance their tactical and strategic objectives. For example, at the time of this writing, the Islamic State in Iraq and Syria (ISIS) has unleashed brutalities on civilians in parts of Syria and Iraq including coercive roundups, forced relocations, mass executions, torture, and beheadings. These actions often are on public display, which serves to indoctrinate and/or intimidate locals into supporting, or not resisting, ISIS's objectives of conquering territory and looting assets. Moreover, such brutalities serve to demean those designated by the atrocity architects as the out-group, which can facilitate the recruitment of people willing to perpetrate genocidal acts and even turn "ordinary people" into participants in genocide. Such elements are observed in virtually all genocides including Rwanda

(1994), Bosnia (1992-1995), Cambodia (1975-1979), and the Holocaust (1933-1945), to name just a few.¹

Those wishing to instill peace also seek to engineer social outcomes. Regarding ISIS advances in summer 2014 and early 2015, for example, Iraqi and Kurdish soldiers (among others) fought ISIS on the ground while the U.S., Jordan, Iran, and other states attacked ISIS from the air with such stated objectives as peace enforcement and protection of vulnerable civilian groups (e.g., the Yazidis). Furthermore, Muslim leaders denounced ISIS and pledged to dissuade Muslim youth from joining the organization.²

Strategic interactions between genocidal architects and potential third-party violence preventers occur in the context of a malleable social context in which the people within such systems also act according to their interests. Given the metaphor between animal breeding and social system breeding, many of the mathematical tools and concepts of evolutionary theory can be adapted to the study of the social evolution of violence and peace. In this article, I take selected aspects of evolutionary game theory to model how genocide architects can artificially manipulate a social environment so that it advances their objectives of isolating and destroying people from an out-group. I also consider how those seeking peace can apply their own techniques to prevent the onset and spread of genocide.

The objective of this article is *not* to do what is often done in research on evolutionary game theory, namely, to prove the existence of evolutionarily stable strategies or states under various conditions within a new model. Rather, I apply a

standard model of evolutionary game theory, in which evolutionary stability conditions are already well known, to a topic in which there has been little application, specifically, the social evolution of genocide across time and geographic space. I consider how genocide architects and third-party genocide suppressors can attempt to manipulate parameters in the standard model to change social environments in ways that are favorable to them. My focus is on social evolution *within the in-group*. Most formal models and many case studies of genocide emphasize interactions between a genocidal in-group and a victim group (and third-party interveners, if present), which represents *inter-group* behavior. Such an emphasis is, of course, important. But equally important, and perhaps more so in some cases, are *social dynamics within an in-group* that cause the in-group to become acceptant of genocide. This article applies evolutionary game theory tools to better understand the emergence of a preference for genocide within an in-group. I also consider how policy interventions can either prevent such a preference from emerging in the first place or how a genocidal preference that has socially evolved can be thwarted.

The article is organized as follows. I offer a brief literature review of evolutionary game theory (first section) and then use a standard evolutionary game theory model to identify characteristics that make an in-group vulnerable to the social evolution of a genocidal preference against an out-group (second section). Under certain conditions, genocide architects can create the social metamorphosis of a peaceful in-group into one that accepts, or does not resist, aggression against an out-group. I refer to this as the social evolution of aggression. Next, I extend the analysis into geographic space by modeling interactions among peaceful and aggressive phenotypes in a geographic area known as a Moore neighborhood (third section). In both the second and third sections I analyze policy interventions in which the social evolution of aggression never gets started or comes to a halt if already underway. I then present two brief cases to illustrate key concepts: the Holocaust (1933-1945) and Côte d'Ivoire (2011) (fourth section). The Holocaust is a "positive" case of genocide in that it occurred, while Côte d'Ivoire is a "negative" case, i.e., one in which genocide seemed likely but did not occur. I conclude with a summary of future research possibilities, key ideas regarding the social evolution of genocide, and policy prescriptions for preventing its onset and spread.³

Brief literature review

Evolutionary game theory models are available in most game theory textbooks and in biology books that contain mathematical methods (e.g., Harrington, 2008; Dixit, *et al.*,

This article applies evolutionary game theory tools to better understand the emergence of a preference for genocide within an in-group. It also considers how policy interventions can either prevent such a preference from emerging in the first place or how a genocidal preference that has socially evolved can be thwarted. Models are applied to the cases of the Holocaust and, in terms of prevention, to Côte d'Ivoire.

2013; Nowak, 2006). Literature on evolutionary game theory and social networks is vast and has been applied in many contexts including economic development (Munshi, 2014), financial contagion (Jackson, 2014), spread of obesity, depression, and drug use (Christakis and Fowler, 2012), social media (Harrigan, *et al.*, 2012; Hodas and Lenman, 2014), and social behavior over geographic space (Ellison, 2000; Galeotti, *et al.*, 2010; Hauert, 2002; Killingback and Doebeli, 1996; Morris, 2000; Ille, 2014). In the terrorism literature, Arce and Sandler (2003; 2009) use evolutionary game theory to model social interactions between fundamentalists and nonfundamentalists within a population.

Many genocide case studies describe the onset and spread of atrocity in terms of "evolution" or "social engineering"; some scholars even note such ideas in their titles (e.g., Messall, 2000; Smith, 2009; Vági, *et al.*, 2013; Verwimp, 2011). Nevertheless, formal evolutionary game theory applications to genocide are rare. In one such application, Gangopadhyay (in Anderton and Brauer, 2016) models the contagion (or acceptance) of genocide among ordinary people. In another, Anderton (2010) models the social evolution of genocide-supporting and genocide-resisting phenotypes in a society populated by hardliners, bystanders, and resisters. This article extends Anderton (2010) here by considering a wider range of parameter possibilities in the space-less model (but with two rather than three phenotypes) and by considering the social evolution of genocide and peace across geographic space.

An evolutionary game theory model of the social evolution of aggression

For the analyses that follow, it is important to understand that a key focus of the article is on the social evolution of genocide acceptance (aggression) or resistance (peace) for people within the in-group. The people from the in-group constitute what I call the population. The potential victim group is the out-group, which is treated as given and is not part of the population. For understanding genocide risk and prevention, group dynamics within the in-group between aggressive and peaceful people-types is potentially just as important as, and at times maybe even more important than, group dynamics between the

in-group and the out-group. Hence, the modeling exercises that follow focus on dynamic and spatial aspects of the social evolution of people-types within the in-group and, specifically, on the “drama” of whether the in-group will tend to become acceptant of genocide against the out-group or whether it will tend to resist such atrocity.

I take as given a regime’s objective to target a people-group for destruction. Except for weapons of mass destruction, atrocity entrepreneurs cannot accomplish the large-scale destruction of an out-group on their own. Hence, they need to enlist a relatively large number of “willing executioners,” which in turn requires social settings wherein such recruitment can succeed. The focus of the model that follows is upon such social settings. Using a standard evolutionary game theory model, I identify conditions that give rise to atrocity-supporting social environments and show how the genocide architects can attempt to foster such communities.⁴

Stage game and fitness equations

Imagine a village with a population of 10,000 people. Assume that an authority group seeks to perpetrate atrocity against an out-group and would like the village to be supportive of its aggression. Assume the potentially targeted out-group encompasses people who are not part of the 10,000 in the village. Some villagers may already have a latent desire for aggression against the out-group; say, there are 1,000 such people (10 percent). Assume the other 9,000 (90 percent) would resist aggression against the out-group under current social conditions. As the village is characterized now, it would not be useful for atrocity architects to perpetrate aggression against the out-group. But can the authority group engineer a social metamorphosis of the village such that it becomes supportive of, or at least not resistant to, their aims? Under certain conditions, the answer is “yes.”

Assume each person in the village has one of two dispositions toward people from the out-group: peaceful (*P*) or aggressive (*A*). The villagers are not genetically hardwired to one disposition or the other; people have free will, so each is free to choose the trait that s/he prefers. Humans are social creatures, so there is a lot of interaction among the people in the village day-by-day. People work with others, buy and sell in the marketplace with others, attend charitable causes and political rallies with others, and recreate with others (e.g., sporting events, picnics). Assume the matrix in Figure 1 represents the payoffs to any two individuals, say Bob (the row player) and Sally (the column player), when they socially interact in the village. Assume Bob and Sally each choose the aggressive trait (*A*) and they socially interact. Based on Figure 1, their interaction leads to payoff *a* for each person. If both

		Sally	
		Aggression (<i>A</i>)	Peacefulness (<i>P</i>)
Bob	Aggression (<i>A</i>)	<i>a, a</i>	<i>b, c</i>
Peacefulness (<i>P</i>)	<i>c, b</i>	<i>d, d</i>	

Figure 1: Payoffs in pairwise encounters in the evolutionary game theory model

display the peaceful (*P*) trait, each would receive payoff *d* from their encounter. If Bob is peaceful and Sally aggressive, Bob would receive *c* and Sally *b* as shown in the lower left cell of the matrix. The obverse interaction is shown in the upper right cell in which Bob receives *b* and Sally *c*.

In evolutionary game theory, Figure 1 is the stage game. It is the matrix that governs the payoffs to individuals in all pairwise social interactions in the village. The payoffs reflect the institutions, history, culture, language, and interpersonal norms of people in the village. Although highly simplified, the matrix in Figure 1 is a type of social genome for the village. Other villages will have different social genomes, and more complex genomes can be represented by more complex stage games. The village’s stage game will determine its traits or phenotypes, i.e., the number of people that are aggressive (*A*) and peaceful (*P*) toward an out-group. Moreover, the atrocity entrepreneurs can attempt to engineer the traits of people in the village so that a greater number adopt aggressiveness. In short, the stage game is amenable to manipulation.

In social evolutionary terms, the traits of people—aggressive (*A*) and peaceful (*P*)—can be thought of as strategies that they display in their social interactions with others. Assume for simplicity that people cannot play a mixed strategy in which they choose *A* or *P* with a random device. Assume two people are randomly drawn from the village and paired with one another. The pairing represents a social encounter in the village and the payoff to each individual in the pair is governed by Figure 1. Many such social encounters occur in the village throughout the day and one can construct the expected payoff for each strategy from such an encounter. Let *N* represent the number of people in the village (say *N*=10,000) and *n_A* the number who initially choose the aggressive trait (say *n_A*=1,000). It follows that the number of people in the village initially choosing peacefulness, *n_P*, is equal to *N*-*n_A*, so that *n_P*=9,000. Given these initial conditions and assuming an individual can be randomly paired with itself (self-play), the probability that a randomly drawn individual from the village would be paired with an aggressive type is

(n_A/N) and with a peaceful type it is $[(N-n_A)/N] \equiv (n_P/N)$.⁵

Based on Figure 1 and the foregoing assumptions, the expected payoff to a villager who adopts the aggressive strategy, F_A , in a random pairwise social encounter is:

$$(1) F_A = a\left(\frac{n_A}{N}\right) + b\left(\frac{N-n_A}{N}\right) = a(r_A) + b(1-r_A),$$

where r_A is the ratio of the number of A -types to the total number in the village ($r_A \equiv n_A/N$). The expected payoff to a villager adopting peace, F_P , in a random pairwise social encounter is:

$$(2) F_P = c\left(\frac{n_A}{N}\right) + d\left(\frac{N-n_A}{N}\right) = c(r_A) + d(1-r_A).$$

In evolutionary game theory, F_A and F_P are the respective fitness of the A and P traits in the village. If aggressiveness is rewarded through, say, advancement in the atrocity leaders' organization, self-preservation, family preservation, or loot, then the A trait would be relatively fit. If, however, aggressiveness is shunned and the peaceful trait is held in high esteem, then the P strategy would be relatively fit. Equations (1) and (2) show that each strategy's fitness depends on the proportion of the villagers playing each strategy and the payoffs generated from the various pairwise social encounters. The average fitness in the village, \bar{F} , is:

$$(3) \bar{F} = F_A\left(\frac{n_A}{N}\right) + F_P\left(\frac{N-n_A}{N}\right) = F_A(r_A) + F_P(1-r_A),$$

where F_i ($i=A,P$) are given in equations (1) and (2).

Replicator dynamics

In social environments, humans look around and learn from the behavior of others. They observe which traits are rewarded and which are penalized, and they tend to choose or mimic traits that are successful. In certain atrocity contexts, people who refuse to support aggression toward an out-group (and thus adopt peacefulness in our model) can be subject to dire penalties including incarceration or execution. In this way, such people are "weeded out" of the population.⁶ To save their lives or careers or the lives of family members and friends, others may act "as if" they support aggression. In the model, I treat such people as adopting the aggressive trait.⁷ It is important to keep in mind in the analyses that follow that the fitness of traits, specifically the incentives people have to adopt

fitter traits and shun less fit traits, drives social evolution in the village.

The replicator or selection dynamics in evolutionary game theory are equations that describe how fitter strategies are adopted and less fit strategies fall out of favor over time. Following Nowak, the selection dynamics in the village is governed by the following difference equations:

$$(4) r_A^{t+1} - r_A^t = r_A^t(F_A - \bar{F}) \text{ and}$$

$$(5) r_P^{t+1} - r_P^t = r_P^t(F_P - \bar{F}),$$

where r_i^j is the ratio of i -types ($i=A, P$) in the village at time j ($j=t+1, t$) and F_A, F_P , and \bar{F} are given by equations (1) to (3), respectively.⁸ Note in equation (4) that if $F_A > \bar{F}$, aggression (A) will be fitter than average and thus fitter than peacefulness (P). Moreover, if $F_A > \bar{F}$ it necessarily follows that peacefulness will be less fit than average (i.e., $F_P < \bar{F}$ in equation 5) and thus less fit than aggression. As such, the number and ratio of people adopting aggression will rise over time and the number and ratio displaying peacefulness will correspondingly decline. It is this dynamic the atrocity entrepreneurs wish to generate. To complete the formal model, let the initial ratio of villagers adopting aggression be r_A^0 and peacefulness $r_P^0 \equiv (1 - r_A^0)$. In the numerical example given earlier, the initial number of aggressors from the $N=10,000$ population was $n_A^0 = 1,000$ and $n_P^0 = 9,000$; hence, the initial ratios are $r_A^0 = 0.1$ and $r_P^0 = 0.9$.

Based on the payoff values a, b, c , and d in the stage game in Figure 1, and assuming no payoff ties for a and c as well as for b and d , there then are four possible cases in which the ratio of aggressive types, r_A , is determined as summarized in Table 1.⁹ In Case 1, the payoffs to aggressive types are low relative to the corresponding payoffs to peaceful types: $a < c$ and $b < d$. The case depicts a simplified social environment in which the village's history, culture, religion, and so forth imply that aggressiveness is unrewarded relative to peacefulness in pairwise social encounters. For Case 2, $a < c$ and $b > d$. Now the payoff for an aggressive type in a "cross encounter" with a peaceful type is more rewarding to the aggressor relative to what s/he would have achieved had s/he been peaceful ($b > d$). In Case 3, $a > c$ and $b < d$. Relative to Case 1, the payoff to an aggressive type in a "same encounter" with an aggressive type is more rewarding to the aggressor relative to what s/he would have achieved had s/he been peaceful ($a > c$). Finally, in Case 4, $a > c$ and $b > d$, the aggressive trait is rewarded relative to peacefulness in pairwise encounters. Such a society would be one in which compliance with the atrocity entrepreneurs would

lead to survival, career advancement, and/or material rewards while noncompliance would correspond to the absence of such benefits.¹⁰

A numerical example: Engineering the social evolution of aggression

As a numerical example of how an atrocity entrepreneur can engineer the social acceptance of aggression against an out-group, assume parameter values consistent with a Case 3 bi-stable outcome: $a=2$, $b=1.5$, $c=1$, $d=2$, and $r_A^0=0.20$ (20 percent). Panel (A) in Figure 2 plots the fitness equations (1) and (2) for these parameters and shows that $r_A^{critical}=0.33$. Since $r_A^0 < r_A^{critical}$, peacefulness is more rewarding (fitter) than aggression and thus more rewarding than average, which leads more people to choose peace over time. Hence, r_A declines over time in Panel (A) until arriving at an evolutionarily stable state in which all villagers adopt peace ($r_A^*=0$).¹¹ Suppose that from the Figure 2, Panel (A) starting point of $r_A^0=0.20$, the atrocity entrepreneurs now insert enough aggressive types from the outside so that r_A^0 rises to 0.25. Such a policy is insufficient to turn the social evolution away from peace because $r_A^0=0.25$ is less than $r_A^{critical}=0.33$. But suppose this effort is coupled with policies that increase payoffs in aggressive/aggressive encounters and reduce payoffs in peaceful/peaceful encounters. Specifically, assume a rises from 2 to 2.2 and d falls from 2 to 1.8. As shown in Panel (B) of Figure 2, the relatively small 10 percent changes in each of a and d cause $r_A^{critical}$ to fall to 0.20. Coupled with the increase in r_A^0 to 0.25 from the “invasion” of aggressive types, r_A^0 is now greater than $r_A^{critical}$. Hence, in Panel (B), the aggressive trait is more rewarding (fitter) than peace, leading more people to choose aggressiveness over time. As such, r_A increases over time in Panel (B) until arriving at an evolutionarily stable state in which all villagers are aggressive types ($r_A^*=1$).

Policy interventions that foster the social evolution of peace

Third-party suppressors of genocide can attempt to implement policies in Figure 2 that will offset the social evolution of aggression. Beginning in Panel (B) of Figure 2, one such policy would be to insert peacekeepers into the village. A sufficiently large number of peacekeepers would reduce the initial proportion of aggressors below the critical value ($r_A^0 < r_A^{critical}$). By becoming part of the peaceful subpopulation of the village, the peacekeepers could tip the village toward the social evolution of peace. But suppose that owing to budget constraints and political realities only a few peacekeepers, or none at all, are inserted into the village. Another policy channel is to manipulate the various payoffs from social encounters in the village in Figure 1 (i.e., the a , b , c , and d values) such that

Table 1: Four possible cases of the social evolution of aggression

- Case 1: $a < c$ and $b < d \Rightarrow$ Peace dominates**
 Parameters imply aggressiveness unrewarded relative to peacefulness in social encounters
- Case 2: $a < c$ and $b > d \Rightarrow$ Peace and Aggression coexist**
 Parameters imply aggressors relatively well rewarded in social encounters with peaceful types
- Case 3: $a > c$ and $b < d \Rightarrow$ Peace and Aggression bistable**
 Parameters imply aggressors relatively well rewarded in social encounters with other aggressors
- Case 4: $a > c$ and $b > d \Rightarrow$ Aggression dominates**
 Parameters imply aggressiveness rewarded relative to peacefulness in social encounters

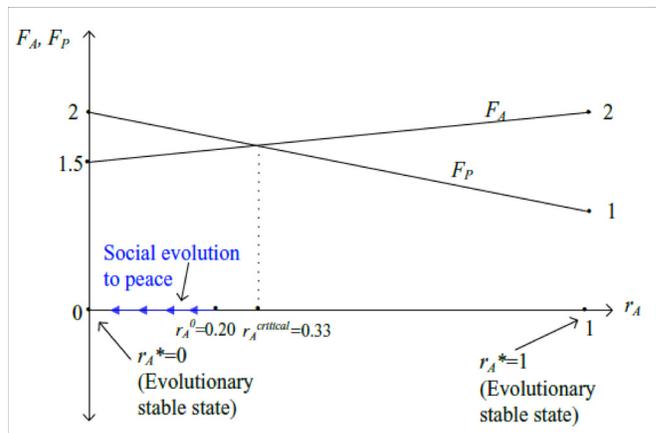


Figure 2(A): Initially $r_A^0=0.20 < r_A^{critical}=0.33 \Rightarrow$ peacefulness is fitter than aggression ($F_P > F_A$) \Rightarrow social evolution over time to evolutionarily stable state at $r_A^*=0$, which is complete peacefulness.

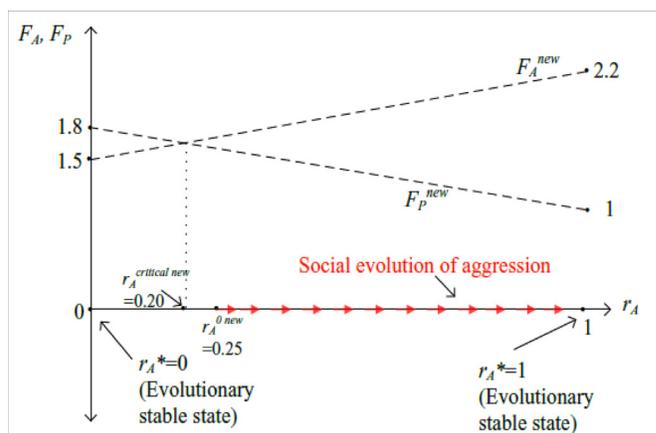


Figure 2(B): An increase in r_A^0 to 0.25, coupled with an increase in a to 2.2 and a decrease in d to 1.8 cause $r_A^{critical}$ to fall to 0.20. Since $0.25 > 0.20$, aggression is fitter than peacefulness ($F_A > F_P$) \Rightarrow social evolution over time to evolutionarily stable state at $r_A^*=1$, which is complete aggression.

the critical value rises to the point that $r_A^0 < r_A^{critical}$. For example, by reducing payoffs to aggressors in social encounters (the a and b payoffs), the F_A^{new} line in Panel (B) of Figure 2 shifts down, which represents a lower expected payoff to the aggressive phenotype for any given r_A . Furthermore, to the extent that peaceful types in the village can achieve greater rewards in social encounters (higher c and d payoffs), the F_P^{new} line in Panel (B) of Figure 2 shifts up, which represents greater expected payoff to the peaceful phenotype. Such line shifts serve to increase the value of $r_A^{critical}$ which, if large enough, will achieve $r_A^0 < r_A^{critical}$ and thus the social evolution of peace. Numerous policies could alter payoffs from interpersonal encounters in the village toward peace, including video surveillance designed to shine a light on those fostering aggression and protect those working for peace, widespread presence of journalists, well-trained police forces, truly robust and expanding safe havens, embassy sheltering for at-risk proponents of peace, threatened or actual sanctions against aggressive types, counter-propaganda through media channels, threats of prosecution against aggressors, and so on.

The model represented in Figure 2 suggests many possible actions that could foster peace, but it also points to a fundamentally important principle regarding the timing of such policies. To the extent that anti-aggression policies are implemented with a lag (policy “dithers”), any given policy efforts become increasingly ineffective. If the social evolution of aggression in Panel (B) of Figure 2 has been going on for some time, then the proportion of people in the village of the aggressive type, r_A , will be relatively high and rising toward $r_A=1$. At this point, bringing in peacekeepers and/or implementing policies to alter the payoffs from social encounters will need to occur at substantially higher levels than otherwise to achieve $r_A < r_A^{critical}$ because r_A is so high. Thus villages can hang on a “knife’s edge” where they are about to tip into a social evolution that will support aggression. On such a knife’s edge, relatively small policy efforts can potentially make a substantial difference for peace, whereas delayed policies will risk becoming increasingly unlikely to alter the aggressive dynamic.¹²

A second, perhaps obvious, fundamental policy principle is implied by Figure 2 and more generally by the analyses in this article: While genocide prevention policy must pay careful attention to *intergroup* dynamics, i.e., protecting the vulnerable out-group from the potentially aggressive in-group, *intragroup* dynamics may be an equally or even more important domain for genocide prevention. By the time the in-group has become acceptant of genocide, prevention policy requires the political will and realization of third-party support to protect the vulnerable out-group from such a dangerous preference.

History has repeatedly shown the fragility of such policy support, with Rwanda 1994 serving as a classic, but certainly not the only, example of such failure. To the extent that third-party support can detect the warning signs of an impending social evolution of aggression and help to prevent the emergence of such a preference within the in-group, a formidable source of genocide can be prevented.

The social evolution of aggression and peace over geographic space

In this section I extend the evolutionary game model of the previous section into geographic space. The introduction of geographic space into the analysis shows, among other things, how the social evolution of aggression and the effectiveness of third-party efforts to suppress aggression are sensitive to population density and other geographic considerations.

Payoffs from social encounters in a Moore neighborhood

Following Nowak, assume each individual in the village is situated on a two-dimensional square lattice in which s/he socially interacts with each immediate neighbor.¹³ In each time period, each individual receives a payoff based on the interactions with the various neighbors. Moreover, each individual observes the payoffs accruing to each neighbor and then adopts in the next period the trait or strategy yielding the highest payoff in the neighborhood. Figure 3 is an example of such characteristics and is known as a Moore neighborhood. In Panel (A) of Figure 3, assume Sally (in the middle of the lattice) initially selects the peacefulness trait (P) and her eight immediate neighbors in the lattice also choose peacefulness. According to the stage game in Figure 1, the payoff to a peaceful individual in a social encounter with another peaceful type is d . In Panel (A) of Figure 3, Sally socially interacts with eight peaceful types in her neighborhood, so she receives a payoff of $8d$. Assuming there is no payoff higher than $8d$ in Sally’s neighborhood, she will choose peacefulness in the next period; otherwise she will switch to the aggressive trait. Panel (B) of Figure 3 presents a different situation for Sally. She is now surrounded by five aggressive and three peaceful neighbors. Based on the stage game of Figure 1, Sally receives a payoff of c from each of the five encounters with an aggressive neighbor and d from each of the three encounters with a peaceful neighbor. Hence, her payoff from the eight encounters in Panel (B) is $5c+3d$. In a case in which c and d are low relative to a and b , respectively, some of Sally’s aggressive-displaying neighbors will have a greater payoff than does Sally. In the next round, Sally thus will have an incentive to switch to the aggressive trait (perhaps for sake of self-preservation, career advancement, etc.).¹⁴

Social evolution of aggression in a Moore neighborhood

I present several dynamic and spatial simulations of the social evolution of aggression to show how an atrocity entrepreneur can engineer outcomes that promote the social acceptance of aggression toward an out-group. I begin with parameter values that align with Case 2 in Table 1 (i.e., $a < c$ and $b > d$) in which there was a coexistence outcome. Although I tether the spatial model here to parameters associated with the space-less model of the previous section, Moore social encounters and payoff generations are not the same as the random pairwise encounters of the space-less model. Hence, the results of the space-less and Moore models will differ.

Assume $a=1$, $c=2$, $b=1.5$, and $d=1$ in the stage game of Figure 1 and the initial proportion of aggressors is $r_A^0=0.20$. Assume the village is made up of $N=10,000$ people. Based on simulation techniques available from EvoLudo, Panel (A) of Figure 4 shows a 100×100 square lattice in which 20 percent of the village's inhabitants are initially aggressive types (shown by the red cells) and 80 percent are peaceful types (blue) [see evoludo.org]. The initial distribution of types is randomly assigned across space by EvoLudo. Panel (B) of Figure 4 shows the result by period 25 of the social evolution of phenotypes in the Moore village. The blue cells represent peaceful types who were peaceful in the previous period and the red cells aggressive types who were aggressive in the previous period. The green cells show peaceful types who were aggressive in the previous period. The yellow cells show aggressive types who were peaceful in the previous period. Figure 4 shows the social evolution to a coexistence outcome in which the proportion of people adopting the aggressive trait will oscillate between about 27 and 41 percent even out to thousands of periods. Particular individuals in the village will switch their strategies over time, depending on relative fitness in sub-neighborhoods, but the proportionate outcome remains in the range of about 27 to 41 percent aggressiveness.

Assume now that $a=2$, $c=1$, $b=1$, and $d=2$ in the stage game of Figure 1 and the initial proportion of aggressors is $r_A^0=0.20$. These parameter values correspond to Case 3 in Table 1 and Figure 2 in the space-less model in which there was a bi-stable outcome. Panel (A) of the previous figure (Figure 4) represents EvoLudo's random distribution of the initial 20 percent aggressive types in the village. Panel (A) of Figure A1, which is placed in the appendix, shows a quick social evolution to peace in the village by period 2 under our new parameters ($r_A^*=0$). Starting from Panel (A), assume now that the atrocity entrepreneurs can increase the payoff in aggressive/aggressive encounters by 25 percent from $a=2$ to $a=2.5$. Figure A1, Panel (B) shows the implications of the parameter change in the village at period 75. Permanent blocks (sub-neighborhoods) of

	N1	N8	N7	
	N2	Sally's payoff = $8d$	N6	
	N3	N4	N5	

Figure 3(A): Sally chooses peacefulness and has 8 peaceful neighbors (N1 to N8; shaded blue) => Sally's payoff = $8d$

	N1	N8	N7	
	N2	Sally's payoff = $5c+3d$	N6	
	N3	N4	N5	

Figure 3(B): Sally chooses peacefulness and has 5 aggressive neighbors (N1 to N5; shaded red) and 3 peaceful neighbors => Sally's payoff = $5c + 3d$

Note: When choosing peacefulness, Sally's payoff from each peaceful neighbor is d and from each aggressive neighbor is c based on the stage game in Figure 1.

aggressive types emerge in the village encompassing about 2.8 percent of the population. This may seem relatively benign, but the village is close to a tipping point in which it could evolve to complete aggression. To illustrate, Panel (C) of Figure A1 retains $c=1$, $b=1$, $d=2$, and $r_A^0=0.20$ but increases parameter a to 2.7. By period 10, aggression has seriously metastasized in the village. By period 23, the village has socially evolved to

complete aggression (not shown).¹⁵

Manipulating a is not the only tool available to atrocity entrepreneurs; parameters b , c , and d can also be manipulated. Assume now that a is back to the value of 2.5 in which blocks of aggressive types occur for 2.8 percent of the population as in Panel (B) of Figure A1. Let b increase by 20 percent from 1 to 1.2 and c fall by 20 percent from 1 to 0.8. Recall that parameter b is the payoff to an aggressive type in a cross encounter with a peaceful type and c is the payoff to a peaceful type in a cross encounter with an aggressive type. These parameter changes tilt the game in favor of aggressors and against peaceful types in cross encounters, everything else the same. Begin from EvoLudo's random distribution of the initial 20 percent aggressive types in the village, such as Panel (A) in Figure 4. In Panel (D) of Figure A1, many concentrations of aggressive types socially evolve by period 5. By period 18, the whole village evolves to the aggressive phenotype such that $r_A^* = 1$ (not shown). Figure A1 shows how relatively easy it is for atrocity entrepreneurs to put the village on the path toward the social evolution of complete aggression. All that was required was a 35 percent favorable change in payoffs to aggressive types in same encounters (increase a from 2 to 2.7) or a 25 percent favorable change in same encounters (increase a from 2 to 2.5) coupled with a 20 percent favorable change in cross encounters (increase b from 1 to 1.2 and decrease c from 1 to 0.8).

Policy interventions for peace in a Moore neighborhood

In the space-less model in the previous section, peacemakers would like to manipulate the village's social genome so that the peaceful trait is rewarded and aggressiveness is unrewarded in pairwise encounters. Hence, if parameters a and b are relatively low and c and d relatively high, Case 1 emerges in Table 1 such that peace dominates everywhere in the village. For less sanguine parameter values, for example those corresponding to the Case 3 bi-stable outcome in Table 1 and Figure 2, the initial number of peaceful types in the village can be critical to insuring that peace rather than aggression evolves. Hence, timely insertion of peacekeepers can be decisive in insuring peace under certain conditions. When considering geographic space in evolutionary dynamics, additional issues come to the fore for peacekeeping consideration, to which I now turn.

Assume $a=2.7$, $b=1$, $c=1$, $d=2$, and $r_A^0=0.10$ in a Moore village. Figure A2, Panel (A), also in the appendix, shows a metastasizing lump of aggressive types by period 3. Left unchecked the lump will grow until the whole village is of the aggressive type (not shown). In Panel (B), 15 peacekeepers are inserted into the lump. Panel (C) shows that the peacekeeping

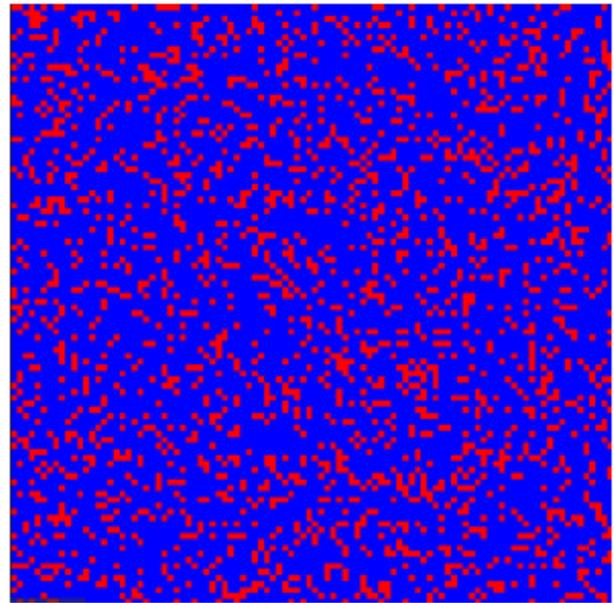


Figure 4(A): Social evolution in the Moore village (for $a=1$, $b=1.5$, $c=2$, $d=1$, $r_A^0=0.20$, $N=10,000$). Time period 0, in which 20 percent of village is aggressive type.

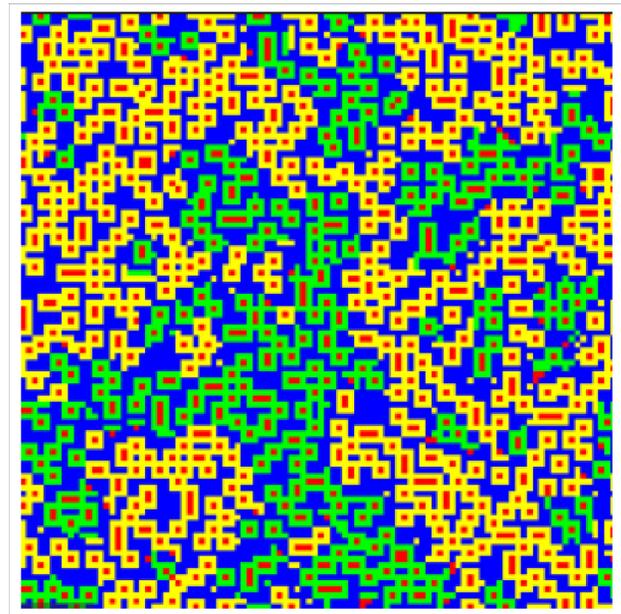


Figure 4(B): Social evolution in the Moore village (for $a=1$, $b=1.5$, $c=2$, $d=1$, $r_A^0=0.20$, $N=10,000$). Time period 25, in which about 40 percent of the village is aggressive type.

effort is too little, too late. By period 15, the lump aggressively metastasizes and, left unchecked, will evolve until the whole village is aggressive (not shown). Panel (D) shows at period 3 a lump that is the same size as that in Panel (C). Similar to the

lump in Panel (A), insertion of 15 peacekeepers is insufficient to stem the growth of aggression (not shown). Hence, insert 20 peacekeepers into the lump in Panel (E). Panel (F) shows that by period 4 aggression is contained and the village only constitutes peaceful types. What is striking about Figure A2 is how the fate of the village hangs on a “knife’s edge.” With 20 rather than 15 peacekeepers, complete peacefulness is sustained in the village rather than the emergence of complete aggression. Furthermore, the spatial location of the peacekeepers is critical. In Figure A2, the peacekeepers needed to be inserted in a concentrated manner into the lump, otherwise the effort would have been completely ineffective.

Again following Nowak, I now consider the role of “walkers” in creating a “big bang” of peace.¹⁶ A walker is a concentrated unit of 10 peaceful types that operates in a world of aggressive types. In Panel (A) of Figure A3, two walker teams are inserted at time 0 into an otherwise completely aggressive Moore village under parameter assumptions that favor peaceful types in same encounters ($d=1$), aggressive types even more in cross encounters ($b=1.51$), but otherwise are not favorable ($a=0$, $c=0$). The village’s population is $N=4,761$, so the 20 peacekeepers imply $r_A^0=0.996$.¹⁷ In 100 simulations, not presented in this article, in which the initial proportion of randomly distributed aggressors was 90 percent, social evolution always converged to complete aggression. But the “walker intervention policy” in Panel (A) leads to a decidedly different outcome. By period 6 the walker teams are close to connecting as shown in Panel (B). Panel (C) shows the early stage of a “big bang” of peace in period 8. By period 15, the social evolution of peace continues to grow (Panel D). Panel (E) shows that by period 60 peace has become dominant in the village, and by period 150, pockets of aggressive types are relatively small and contained as shown in Panel (F). Similar to Figure A2, the message of Figure A3 is that the insertion of peacekeepers in both timely and geographically concentrated manners is critical for fostering the social evolution of peace.

The Holocaust and Côte d’Ivoire

In this section I apply several key ideas of the evolutionary game theory model to two cases. The first, the Holocaust, is a “positive” case of genocide in that genocide occurred. The case is also notable in that little effort was made by possible third-party suppressors to stop the genocide. The second, Côte d’Ivoire, is a “negative” case of genocide in that the risk of genocide seemed high but full-blown genocide did not occur. This case is also notable owing to relatively substantial and early third-party efforts to suppress genocide.

The Nazi example: Engineering the social evolution of aggression in occupied Europe

The payoffs a , b , c , and d in the stage game in Figure 1 are the outcomes from social interactions in the village, which are rooted in the village’s institutions, culture, and history. The payoffs can be thought of as the village’s heritage or social genomic endowment. I do not model the origins of the payoffs, but take them as given. In villages influenced or controlled by genocide architects, the social genome is subject to social engineering in which violence entrepreneurs attempt to foster a social metamorphosis to aggression. The previous analyses, summarized in Table 1, show that increases in a and b and decreases in c and d move the village from Case 1 in which peace dominates toward Case 4 in which aggression dominates. In moving toward Case 4, the intermediate Case 3 (with high a and low c) could also be amenable to the social evolution of aggression provided a sufficiently large invasion of aggressors came into the village (see Figure 2). Hence, the violence entrepreneurs would like to increase a and b , decrease c and d , and/or insert an outside gang of aggressors into the village. What do these constructs represent in a real-world setting of potential social evolution of aggression? I address this question by presenting selected social evolutionary techniques deployed by Nazi officials in the occupied territories of Europe during world war two.

In the foundational book in the field of genocide studies, *Axis Rule in Occupied Europe*, Raphael Lemkin (1944, pp. xi, 79) described genocide as a “synchronized attack” designed to destroy the existence of a victim group and replace it with the national pattern of the oppressor. Lemkin (1944, pp. xiii-xiv) identified eight dimensions or “fields” of group destruction:

[I]n the political field (by destroying institutions of self-government and imposing a German pattern of administration, and through colonization by Germans); in the social field (by disrupting the social cohesion of the nation involved and killing or removing elements such as the intelligentsia, which provide spiritual leadership ...); in the cultural field (by prohibiting or destroying cultural institutions and cultural activities; by substituting vocational education for education in the liberal arts, in order to prevent humanistic thinking ...); in the economic field (by shifting wealth to Germans and by prohibiting the exercise of trades and occupations by people who do not promote Germanism “without reservations”); in the biological field (by a policy of depopulation and by promoting procreation by Germans in occupied countries); in the field of physical existence (by introducing a starvation rationing system for non-Germans and by mass

killings, mainly Jews, Poles, Slovenes, and Russians); in the religious field (by interfering with the activities of the Church, which in many countries provides not only spiritual but national leadership); in the field of morality (by attempts to create an atmosphere of moral debasement through promoting pornographic publications and motion pictures, and the excessive consumption of alcohol).

What occurred in each of these eight fields was a social advancement of people who supported, or did not resist, the Nazi program and a social degradation of those who did resist. In the context of the simplified stage game in Figure 1, the Nazis implemented policies that increased payoffs to supporters of aggression when aggressive phenotypes interacted (increase a) and when aggressive and peaceful types interacted (increase b). On the flip side, peaceful types were harmed through policies that decreased payoffs to supporters of peace when peaceful and aggressive types interacted (decrease c) and when peaceful types interacted (decrease d). Furthermore, the Nazis implemented invasions of aggressive types into locales (increase r_A^0) to foster the social evolution of aggression. Table 2 presents selected Nazi techniques that promoted the social evolution of aggression categorized according to the parameters of the evolutionary game model.

Côte d'Ivoire 2011: Genocide averted?

Acts of “lower-level” intentional violence against civilians and an elevated risk of mass atrocities emerged in Côte d'Ivoire in the context of the disputed presidential election of 28 November 2010 between incumbent Laurent Gbagbo and challenger Alassane Ouattara. Gbagbo claimed victory in the election, but the U.N. Secretary General and the Economic Community of West African States (ECOWAS) maintained that Ouattara was the winner. In the period during and immediately after the election, the Gbagbo regime instituted a curfew, prevented the dissemination of election results, canceled 660,000 votes for Ouattara, and carried out repressive violence against civilians. By early 2011, reports emerged of patterns of civilian killings across the country and suspected mass graves. According to Human Rights Watch, supporters of the Gbagbo regime carried out executions, rapes, and immolations against northern Ivoirians, West African immigrants, and political supporters of Ouattara. Civilian abuses were also carried out by the pro-Ouattara Republican Forces, although on a comparatively smaller scale. Estimates of civilians intentionally killed in Côte d'Ivoire in 2011 range from several hundred to at least 3,000. On 19 January 2011, the Special Advisors to the Secretary General on the Prevention of Genocide and R2P communicated their grave

concerns about civilian violence in Côte d'Ivoire including a statement that clashes “*if not checked*, could culminate in *mass atrocities*.” In the Holocaust, the Nazi regime’s social engineering of atrocity acceptance by the in-group against Jews and other groups was essentially uncountered by possible third-party suppressors of genocide. In Côte d'Ivoire, however, attempts by Gbagbo and his supporters to retain political power by whatever means necessary, including atrocity, were met with significant counter-efforts by the U.N., ECOWAS, and other organizations.¹⁸

In the context of the evolutionary game model developed in this article, I treat the Gbagbo regime of leaders, civil servants, and supporters as the in-group over which an intragroup social evolution of aggression or peace will play out. If this in-group socially evolves to acceptance of aggression, then the risk of atrocities against those designated as outsiders will be high. In the context of the model, the payoffs to those in the in-group supporting the aggressive stance of the Gbagbo regime could be high relative to those resisting owing to (1) serious ramifications from not going along with those in power and/or (2) potentially harsh reprisals for Gbagbo supporters given previous repression should power be lost. This suggests in Figure 1 and in Table 1, $a > c$ and $b > d$ such that aggression dominates peace. If not checked, such a setting would be one in which atrocities against out-groups would become socially acceptable to those within the in-group.

The checks, which I characterize as third-party efforts to foster the social evolution of peace in the model, were several and ultimately significant in reducing the risk of mass atrocity. For example, on 20 December 2010, U.N. Security Council Resolution 1962 was unanimously adopted in which Ouattara was recognized as the winner of the election. The resolution also extended the United Nations Operation in Côte d'Ivoire (UNOCI) including requests to strengthen the mission with deployment of Ukrainian helicopters. Meanwhile, the African Union, ECOWAS, France, and the U.S. recognized Ouattara as the president-elect. These actions, alone, were insufficient to lead Gbagbo to step down, but additional policies followed. For example, ECOWAS “moved to suspend the government’s access to banking services, undercutting Gbagbo’s ability to pay civil servants and soldiers and thereby reducing his allies to a small coterie of dedicated followers.” In the context of the evolutionary game model, such actions reduced the payoffs to supporters of aggression against out-groups leading to reductions in a and b in Figure 1 and Table 1. This, in turn, tilts the evolutionary outcome away from the dominance of aggression to improved prospects for peace. Perhaps most significant in reducing the expected payoff to aggression occurred when the Gbagbo regime appeared ready to deploy

heavy weapons against civilians in early April 2011. The threat was countered by U.N. peacekeepers and the French *Force Licorne*, which attacked Gbagbo's military assets in Abidjan. Gbagbo's arrest followed on 11 April 2011.¹⁹

Conclusions

People-groups are amenable to forms of social engineering in which their interpersonal interactions and population mixes are manipulated to support the aims of a small group of social engineers. In this article, the social engineers are the genocide architects who seek to socially transform people-groups away from peacefulness to an aggressive posture toward out-groups. Using evolutionary game theory tools, I analyzed how people socially interact with others to highlight conditions under which atrocity engineers can transform a village into supporting, or not resisting, aggression against an out-group. I also analyzed conditions under which peacemakers could prevent the social evolution of aggression through various policy efforts.

The key policy messages of the article are several. First, under a fairly wide range of conditions, people-groups can socially exist on a "knife's edge" in which small social engineering efforts toward aggression or peace can have dramatic effects. Several of the model simulations showed that a small effort to foster aggression by the atrocity entrepreneurs could tip an otherwise peaceful village into a social metamorphosis in which the whole village comes to support aggression. The tipping can go the other way. Relatively small but well-timed and well-located peacekeeping efforts can neutralize the social evolution of aggression and allow peace to persist in a village. The fate of a village can truly hang in the balance between the social engineering efforts of atrocity entrepreneurs and peacemakers. Second, geographic space matters in understanding the social evolution of aggression. In extended simulations, not presented in this article, I found that

Table 2: Manipulating the social genome: Selected Nazi techniques that foster the social evolution of aggression in the evolutionary game model

1. Increase payoffs to aggressive phenotypes in same encounters (increase a)
Mass changes in law that favor the occupiers and disfavor the occupied (p. 25)
Support of Germanism rewarded in economic life (pp. 195, 206, 224)
2. Increase payoffs to aggressive phenotypes in cross encounters (increase b)
Mass changes in law that favor the occupiers and disfavor the occupied (p. 25)
Appropriation of property from occupied groups (pp. 37-40, 144)
Legal rights of occupied replaced by "grace" of occupant (p. 71)
Support of Germanism rewarded in economic life (pp. 195, 206, 224)
3. Decrease payoffs to peaceful phenotypes in cross encounters (decrease c)
Mass changes in law that favor the occupiers and disfavor the occupied (p. 25)
Appropriation of property from occupied groups (pp. 37-40)
Refuse unemployment relief to those unwilling to work for Germany (p. 69)
Legal rights of occupied replaced by "grace" of occupant (p. 71)
Severe penalties to those who help victim groups (p. 77)
Resistance to Germanism penalized in economic life (pp. 195, 206, 224)
4. Decrease payoffs to peaceful phenotypes in same encounters (decrease d)
Mass changes in law that favor the occupiers and disfavor the occupied (p. 25)
Disrupt centers of political resistance (p. 67)
Separate families (p. 67)
Exclude people from liberal arts education (pp. 84, 229)
Undermine religious affiliations and leadership in occupied territories (p. 89)
Create an atmosphere of moral debasement (pp. 89-90)
Appropriation of the gains from trade within occupied territories (pp. 127, 229)
Imprisonment or liquidation of key leaders (p. 139)
Resistance to Germanism penalized in economic life (pp. 195, 206, 224)
Assembly restricted (p. 231)
5. Increase initial ratio of aggressive phenotype in locales (increase r_A^0)
Ideological penetration and fifth column support of Nazism (pp. 19, 83, 137, 237)
Mass deportations of native populations (pp. 21, 67)
Subsidies and tax breaks for German settlers in occupied territories (pp. 21, 63, 225)
Insertion of "colonization staffs" into occupied territories (p. 21)
Currency manipulation to finance insertion of aggressive types into areas (pp. 51-53)

Notes: Parameters a , b , c , d , and r_A^0 are parameters of the stage game of the basic model in Figure 1. Page numbers in the table refer to pages in Lemkin (1944).

social evolution in the moderately dense Moore village can play out very differently over time and space relative to villages with higher or lower population densities. For example, "walker" peacekeeping interventions in a Moore village can lead to a "big bang" of peace, but the same conditions in a village with lower population density can lead to a muted outcome for peace. I also found that the deterministic space-less model leads to definitive outcomes whereas the outcomes in models with geographic space can be sensitive to the initial random spatial distribution of the population mix. Third, under conditions in which the social evolution of aggression can metastasize within a village, early and geographically well-located interventions for peace are

essential to stop aggressiveness in its tracks. Fourth, any general claims about the ability or inability of peacekeepers to prevent the social evolution of aggression should be looked at with doubt. Social context is immensely important for understanding how people relate or misrelate to one another and such contexts vary widely across the world and within states, cities, towns, rural areas, and neighborhoods. In some social settings, a small peacekeeping effort can have dramatic effects in fostering a more peaceful social context among people from an in-group. In other settings, the aggressive fate of the village may be sealed and scarce peacekeeping resources might be better deployed elsewhere. Finally, it is clear that genocide prevention policy must be concerned with protecting a vulnerable out-group from a potentially aggressive in-group; this is the domain of intergroup dynamics. But perhaps of equal or even greater importance is a policy focus on dynamics within the in-group, for it is here that a social evolution toward aggression can occur and a preference for genocide can arise. Stopping the emergence of such a preference in the first place will be better for a potential victim group, and probably much cheaper too, than trying to thwart such a preference after it has become actualized.

Although the article is not focused on empirical hypothesis testing, an emphasis on evolutionary social dynamics within an in-group suggests bottom-up or micro-oriented risk factors for genocide including certain historical, cultural, psychological, leadership, and village-specific characteristics. The empirical literature on genocide is dominated by work in which the characteristics of states and/or rebel groups drive the formulation of hypotheses and the data sources used. This, of course, is valuable research, but it can be complemented by theoretically informed empirical research on in-group characteristics that foster or inhibit genocide.

Future research on theoretical modeling of the social evolution of aggression and peace should consider neighborhood assumptions beyond Moore. For example, individuals in a village could be treated as triangular (having three direct neighbors), quadrangle (four direct neighbors, which is known as a von Neumann village), pentagonal (having five direct neighbors), and so forth such that a wide range of population densities could be modeled. EvoLudo's simulation techniques already allow very dense neighborhoods of 24, 48, 80, and even higher levels, with results that can vary significantly from Moore geography (8 direct neighbors). Furthermore, there are alternative methods for modeling replicator dynamics including variations in the degree to which payoffs translate into fitness, potential for subgroups of the population to only interact with those they desire, migrations to and from the population (e.g., births, deaths, refugees), and

stochastic elements. In addition, the spatial social evolution of aggression and peace can consider three or more phenotypes and apply such models to genocide prevention. Potentially important geographic elements such as forests, mountain ranges, borders, waterways, safe havens, and urban/rural population density variations in sub-neighborhoods could also be incorporated into such models. Furthermore, the standard evolutionary game theory model opens the door to possible new research including constrained optimization modeling in which a third party seeks to maximize prospects for the social evolution of peace by choosing various policies (e.g., insertion of peacekeepers, policies to reduce a and b , geographic locations of such efforts) subject to a resource and/or time constraint. Additionally, one could use game theory to model the strategic contest between the atrocity entrepreneurs and third parties over the social evolution of the village. Another potentially productive research avenue is to use analytical functions in the stage game of Figure 1 rather than exogenous parameters. Such functions would allow one to explicitly model various elements that determine the payoff values that individuals receive in social encounters such as historical conditions, cultural norms, and social-psychological perceptions of people. For example, explicitly modeling the social-psychological perceptions that people have of themselves and their environment and how these perceptions affect payoffs would allow one to study how people from the in-group can rationalize their support of genocide and how this can reinforce the social evolution of aggression. We see a potentially rich future of policy-relevant interdisciplinary research based on evolutionary game theory in which new insights into the social evolution of genocide and other forms of civilian violence are revealed and the prevention of such atrocities is promoted.²⁰

Notes

I am grateful to Daniel Arce, Jurgen Brauer, J. Paul Dunne, and several anonymous referees for helpful comments on earlier versions of this article. I alone am responsible for any errors and omissions and for the views expressed.

1. Ordinary people: See Waller (2007).
2. BBC (2014); Constable (2014).
3. The 1948 United Nations Convention on the Prevention and Punishment of the Crime of Genocide defines *genocide* as "any of the following acts committed with intent to destroy, in whole or in part, a national, ethnical, racial or religious group, as such: (a) Killing members of the group; (b) Causing serious bodily or mental harm to members of the group; (c) Deliberately inflicting on the group conditions of life calculated to bring about its physical destruction in whole or in

part; (d) Imposing measures intended to prevent births within the group; (e) Forcibly transferring children of the group to another group” (United Nations, 1948). *Mass killing* involves intentional killing of civilians, but the perpetrators do not seek to destroy a group as such (Waller, 2007, p. 14). Other atrocities include *war crimes*, *crimes against humanity*, and *ethnic cleansing*. I use the term *mass atrocity* to encompass genocide, mass killing, and other atrocity crimes carried out at a large scale. For further definitional issues, as well as controversies, see Curthoys and Docker (2008), Moses (2010), and Anderton and Brauer (2015). Although the social evolutionary modeling in this article focuses on genocide, the models and results can be applied to other mass atrocity contexts and to other forms of violence (e.g., nation-state war, civil war, terrorism) in which a social group might come to support, or to resist, those wishing to initiate violence.

4. Willing executioners: Goldhagen (1996).
5. Self-play is an assumption that simplifies the probability calculations for random draws without substantially affecting the results. Specifically, without self-play, the probability that an *A*-type individual would be paired in a random draw with another *A*-type is $[(n_A-1)/(N-1)]$ and with a *P*-type $[(N-n_A)/(N-1)]$. Self-play allows one to treat these probabilities as (n_A/N) and $[(N-n_A)/N]$, respectively.
6. I do not model the weeding out of resisters from the population.
7. A model with more than two phenotypes could treat feigning aggressiveness as a third phenotype, as in Arce and Sandler (2009).
8. Nowak (2006, p. 47).
9. Four possible cases: (Nowak, 2006, p. 50; Weibull, 1995, pp. 40-41). Not allowing for payoff ties for *a* and *c* and for *b* and *d* eliminates special cases that are not essential to the article’s main points.
10. Since $r_A+r_P=1$ at all times, the replicator dynamics in equations (4) and (5) can be represented by one difference equation: $r_A^{t+1} - r_A^t = r_A^t(1-r_A^t)(F_A - \bar{F})$ (see Nowak, 2006, pp. 47-48). Plugging F_A and F_P from equations (1) and (2) into this difference equation leads to $r_A^{t+1} - r_A^t = r_A^t(1-r_A^t)[(a-b-c+d)r_A^t + b-d]$ (see Nowak 2006, p. 50). This equation can be used to derive the four cases in Table 1. For Case 1 ($a < c$ and $b < d$), $r_A^{t+1} - r_A^t < 0$ and r_A falls to zero (peace dominates). For Case 2 ($a < c$ and $b > d$), $r_A^{t+1} - r_A^t$ reaches zero for an interior value of r_A (aggression and peace coexist). In Case 3 ($a > c$ and $b < d$), $r_A^{t+1} - r_A^t$ reaches zero at an interior value of r_A , but it is unstable. From this point onward, a small random increase in r_A will cause social evolution to complete aggression or a small random decrease in r_A will cause social evolution to complete peacefulness (aggression and peace bi-stable). In Case 4 ($a > c$ and $b > d$), $r_A^{t+1} - r_A^t > 0$ and r_A rises to 1 (aggression dominates).

11. According to Maynard Smith (1982, p. 204), a “population is said to be in an ‘evolutionarily stable state’ if its genetic composition is restored by selection after a disturbance, provided the disturbance is not too large. Such a population can be genetically monomorphic [one phenotype] or polymorphic [more than one phenotype].” In our context of social evolution, possible genetic compositions are aggression and peacefulness. In Panel (A) of Figure 2, random generation of a small number of aggressive types (i.e., a small disturbance) would cause r_A to rise slightly above $r_A^*=0$. Since peace would be fitter than aggression following this disturbance, ($F_P > F_A$ at small r_A), selection would reduce the ratio of aggressive types and restore $r_A^*=0$.

12. Although not part of this article, the evolutionary game model can be used to calculate (within the model) the precise number of peacekeepers and/or new social encounter payoff values necessary to initiate the social evolution of peace at any given point on the village’s social evolutionary time-line.

13. Nowak (2006, ch. 9).

14. Although the boundaries of the lattices in Figure 3 would appear to imply that individuals on the edges do not have eight neighbors, simulations of spatial evolution in a Moore neighborhood can wrap such edges around the square to generate a torus. Toroidal geometry implies that each cell in the grid is equivalent to each other cell and each individual has eight neighbors (Nowak, 2006, p. 148).

15. The time period of convergence and, for some parameter sets, the proportion of phenotypes that persist over time, depend on EvoLudo’s random spatial distribution of initial phenotypes.

16. Nowak (2006, pp. 159-160).

17. $N=4,761$ was Nowak’s (2006) assumption in his walker analysis, which I follow here.

18. Reports of mass killings and mass graves: Bellamy and Williams (2011, p. 832). Human Rights Watch: HRW (2011, pp. 4-6). Estimates of civilians killed: Pettersson (2014); Eck and Hultman (2007); HRW (2011). Special Advisors quote: Quoted in Bellamy and Williams (2011, p. 833; my emphases).

19. Helicopters: Bellamy and Williams (2011, p. 832). Quote: Bellamy and Williams (2011, p. 833). *Force Licorne*: HRW (2011, pp. 6-7).

20. More phenotypes and genocide prevention: See, e.g., Killingback and Doebeli (1996). Analytical functions: As just one example, Galeotti, *et al.* (2010) consider payoff functions that exhibit strategic complementarity or strategic substitutability across neighbors, positive and negative externalities across neighbors, public goods technologies among neighbors such as best shot and aggregate effort, incomplete information, etc. Social-psychological perspectives: Such elements include obedience to authority, moral disengagement, blaming people from the out-group, escalating commitments, ritualistic treatment of people from the out-group, diffusion of responsibility, conformity to peer

pressure, repression of conscience, and a reordering of one's being into acceptance of aggression (see Waller, 2007, chs. 6-8).

References

- Anderton, C.H. 2010. "Choosing Genocide: Economic Perspectives on the Disturbing Rationality of Race Murder." *Defence and Peace Economics*. Vol. 21, No. 5-6, pp. 459-486.
<http://dx.doi.org/10.1080/10242694.2010.513478>
- Anderton, C.H. and J. Brauer. 2015. "Genocide," in J.G. Backhaus, general ed., and G. Ramello, section ed., *Encyclopedia of Law and Economics*. Berlin: Springer.
http://dx.doi.org/10.1007/978-1-4614-7883-6_581-1
- Anderton, C.H. and J. Brauer, eds. Forthcoming. *Economic Aspects of Genocides, Other Mass Atrocities, and Their Prevention*. New York: Oxford University Press.
- Arce, D.G. and T. Sandler. 2003. "An Evolutionary Game Approach to Fundamentalism and Conflict." *Journal of Institutional and Theoretical Economics*. Vol. 159, No. 1, pp. 132-154.
<http://dx.doi.org/10.1628/0932456032975096>
- Arce, D.G. and T. Sandler. 2009. "Fitting In: Group Effects and the Evolution of Fundamentalism." *Journal of Policy Modeling*. Vol. 31, No. 5, pp. 739-757.
<http://dx.doi.org/10.1016/j.jpolmod.2008.07.002>
- Bellamy, A.J. and P.D. Williams. 2011. "The New Politics of Protection? Côte d'Ivoire, Libya and the Responsibility to Protect." *International Affairs*. Vol. 87, No. 4, pp. 825-850.
<http://dx.doi.org/10.1111/j.1468-2346.2011.01006.x>
- [BBC] British Broadcasting Corporation. 2014. "UK Imams Condemn Isis in Online Video." *BBC News Online*. 11 July 2014. <http://www.bbc.com/news/uk-28270296> [accessed 16 December 2014].
- Christakis, N.A. and J.H. Fowler. 2012. "Social Contagion Theory: Examining Dynamic Social Networks and Human Behavior." *Statistics in Medicine*. Vol. 32, No. 4, pp. 556-577.
<http://dx.doi.org/10.1002/sim.5408>
- Constable, P. 2014. "U.S. Muslim Leaders Denounce Islamic State, Pledge to Dissuade Youth from Joining." *Washington Post Online*. 10 September 2014. www.washingtonpost.com [accessed 15 December 2014].
- Curthoys, A. and J. Docker. 2008. "Defining Genocide," pp. 9-41 in D. Stone, ed. *The Historiography of Genocide*. New York: Palgrave Macmillan.
- Dixit, A.K., S. Skeath, and D. H. Reiley, Jr. 2013. *Games of Strategy*. 4th ed. New York: W.W. Norton.
- Eck, K. and L. Hultman. 2007. "One-sided Violence against Civilians in War: Insights from New Fatality Data." *Journal of Peace Research*. Vol. 44, No. 2, pp. 233-246.
<http://dx.doi.org/10.1177/0022343307075124>
- Ellison, G. 2000. "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution." *Review of Economic Studies*. Vol. 67, No. 1, pp. 17-45.
<http://dx.doi.org/10.1111/1467-937X.00119>
- Galeotti, A., S. Goyal, M.O. Jackson, F. Vega-Redondo, and L. Yariv. 2010. "Network Games." *Review of Economic Studies*. Vol. 77, No. 1, pp. 218-244.
<http://dx.doi.org/10.1111/j.1467-937X.2009.00570.x>
- Goldhagen, D.J. 1996. *Hitler's Willing Executioners: Ordinary Germans and the Holocaust*. New York: Knopf.
- Harrigan, N., P. Achananuparp, and E. Lim. 2012. "Influentials, Novelty, and Social Contagion: The Viral Power of Average Friends, Close Communities, and Old News." *Social Networks*. Vol. 34, No. 4, pp. 470-480.
- Harrington, J. 2008. *Games, Strategies and Decision Making*. New York: Worth Publishers.
- Hauert, C. 2002. "Effects of Space in 2x2 Games." *International Journal of Bifurcation and Chaos*. Vol. 12, No. 7, pp. 1531-1548.
<http://dx.doi.org/10.1142/S0218127402005273>
- Hodas, N.O. and K. Lerman. 2014. "The Simple Rules of Social Contagion." *Scientific Reports*. Vol. 4, Article No. 4343, pp. 1-7.
- [HRW] Human Rights Watch. 2011. "They Killed Them Like It Was Nothing": The Need for Justice for Côte d'Ivoire's Post-Election Crimes. New York: Human Rights Watch.
- Ille, S. 2014. "The Dynamics of Norms and Conventions under Local International and Imitation." *International Game Theory Review*. Vol. 16, No. 3, Article 1 [23 pages].
- Jackson, M.O. 2014. "Networks in the Understanding of Economic Behaviors." *Journal of Economic Perspectives*. Vol. 28, No. 4, pp. 3-22.
<http://dx.doi.org/10.1257/jep.28.4.3>
- Killingback, T. and M. Doebeli. 1996. "Spatial Evolutionary Game Theory: Hawks and Doves Revisited." *Proceedings: Biological Sciences*. Vol. 263, No. 1374, pp. 1135-1144.
<http://dx.doi.org/10.1098/rspb.1996.0166>
- Lemkin, R. 1944. *Axis Rule in Occupied Europe: Laws of Occupation, Analysis of Government, Proposals for Redress*. Washington, D.C.: Carnegie Endowment for International Peace.
- Maynard Smith, J. 1982. *Evolution and the Theory of Games*. New York: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9780511806292>
- Messall, R. 2000. "The Evolution of Genocide." *Human Life Review*. Vol. 26, No. 1, pp. 47-75.
- Morris, S. 2000. "Contagion." *Review of Economic Studies*. Vol. 67, No. 1, pp. 57-78.
<http://dx.doi.org/10.1111/1467-937X.00121>
- Moses, A.D. 2010. "Raphael Lemkin, Culture, and the Concept of Genocide," pp. 19-41 in D. Bloxham and A.D. Moses, eds. *The Oxford Handbook of Genocide Studies*. New York: Oxford University Press.
<http://dx.doi.org/10.1093/oxfordhb/9780199232116.013.0002>
- Munshi, K. 2014. "Community Networks and the Process of Development." *Journal of Economic Perspectives*. Vol. 28, No. 4, pp. 49-76.
<http://dx.doi.org/10.1257/jep.28.4.49>
- Nowak, M.A. 2006. *Evolutionary Dynamics: Exploring the*

Equations of Life. Cambridge, MA: Harvard University Press.

Pettersson, T. 2014. "UCDP One-sided Violence Codebook, Version 1.4." Uppsala Conflict Data Program, Department of Peace and Conflict Research, Uppsala University.

Smith, E.O. 2009. "Evolution, Primates, and Subaltern Genocide," pp. 159-187 in N.A. Robins and A. Jones, eds. *Genocides by the Oppressed: Subaltern Genocide in Theory and Practice*. Bloomington, IN: Indiana University Press.

Vági, Z., L. Csósz, and G. Kádár. 2013. *The Holocaust in Hungary: Evolution of Genocide*. Lanham, MD: AltaMira Press.

Verwimp, P. 2011. "The 1990-92 Massacres in Rwanda: A Case of Spatial and Social Engineering." *Journal of Agrarian Change*. Vol. 11, No. 3, pp. 396-419. <http://dx.doi.org/10.1111/j.1471-0366.2011.00315.x>

Waller, J. 2007. *Becoming Evil: How Ordinary People Commit Genocide and Mass Killing*. 2nd ed. New York: Oxford University Press.

Weibull, J.W. 1995. *Evolutionary Game Theory*. Cambridge, MA: MIT Press.

Appendix

This Appendix contains Figures A1 (this page), and A2 and A3 (next page).

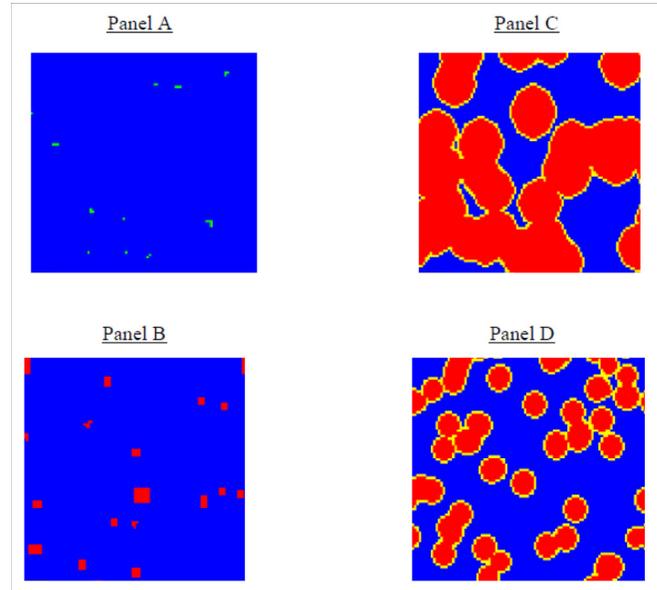


Figure A1: Social evolution in the Moore village for various parameters. In Panel A, social evolution is to complete peacefulness by period 2 for $a=2$, $b=1$, $c=1$, $d=2$, $r_i^0=0.20$, $N=10,000$. In Panel B, the payoff in aggressive/aggressive encounters rises to $a=2.5$, leading to convergence to 2.8% aggressive types in blocks by period 75. In Panel C, the payoff in aggressive/aggressive encounters rises to $a=2.7$, leading to a tipping point for the social evolution of aggression and a substantial proportion of aggressive types by period 10. In Panel D, the social evolution toward aggression is shown at period 5 for $a=2.5$, $b=1.2$, $c=0.8$, $d=2$, $r_i^0=0.20$, $N=10,000$.

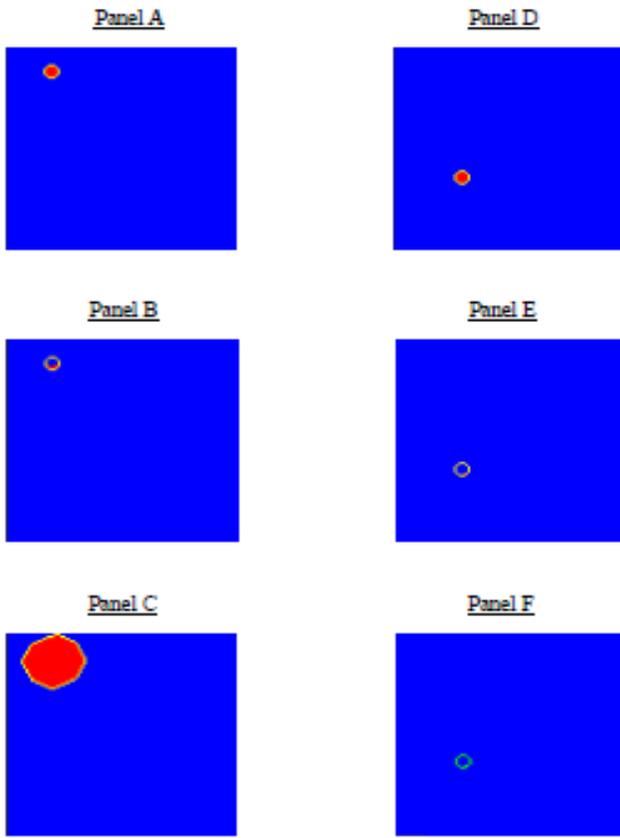


Figure A2: Peacekeeping and the fate of the Moore village for various parameters. In Panel A, a lump of aggressive types emerges by period 3 for $a=2.7$, $b=1$, $c=1$, $d=2$, $r_i^0=0.10$, $N=10,000$. In Panel B, 15 peacekeepers are inserted into the period 3 aggressive lump. The peacekeepers are too little, too late as shown in Panel C by the seriously metastasizing lump at period 15. In Panel D, a period 3 lump of aggressive types emerges similar to Panel A. In Panel E, 20 peacekeepers are inserted into the period 3 aggressive lump. The aggressive lump is contained by period 4 in Panel F.

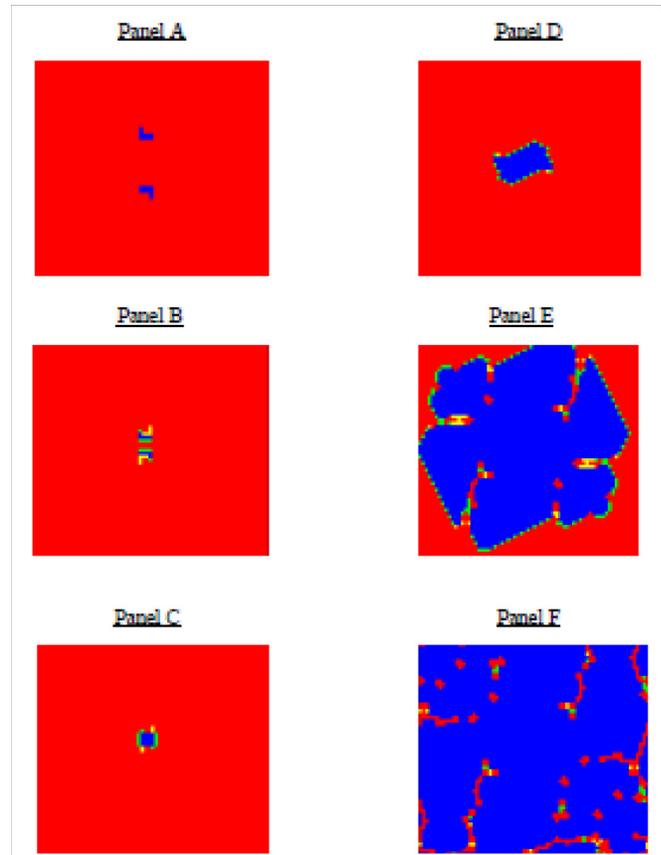


Figure A3: “Walkers” and the “big bang” of peace in the Moore village. In Panel A, 2 walker teams of 10 peaceful types are inserted at time 0 for $a=0$, $b=1.51$, $c=0$, $d=1$, $r_i^0=0.996$, $N=4,761$. In Panel B, the walker teams are close to concentrating by period 6. In Panel C, the walkers connect and the big bang of peace begins at period 8. In Panel D, the big bang of peace continues to grow at period 15. In Panel E, peace is becoming dominant at period 60. In Panel F, the village is made up of mostly peaceful types by period 150. *Source:* Adapted from Nowak (2006).