MODELING IDENTITY-BASED CONFLICT AND GENOCIDE – AN APPROACH INFORMED BY COMPLEXITY THEORY AND COMPUTATIONAL SOCIAL SCIENCE

by

Elizabeth Munira von Briesen

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Computing and Informatics

Charlotte

2020

Approved by:

____

Dr. Mirsad Hadžikadić

____

Dr. Samira Shaikh

____

Dr. John Cox

____

Dr. Amy Canevello

____

Dr. Min Shin

____

Dr. William Rand

ABSTRACT

ELIZABETH MUNIRA VON BRIESEN. Modeling identity-based conflict and genocide – an approach informed by complexity theory and computational social science. (Under the direction of DR. MIRSAD HADŽIKADIĆ)

This research designs the Eris model, a novel, agent-based computational model that can aid researchers in answering questions regarding the explanation and prediction of identity-based conflict and genocide. At its most fundamental level, this work draws on complexity theory and research from the domain of computational social science in order to determine the most effective approach. Recent models of civil violence and violence around the issue of identity have taken advantage of advances in computing power to model both theoretical and highly realistic scenarios, validating on real-world events in some cases, and providing "artificial worlds" in which to perform scenario analysis. The Eris model implements a novel combination of micro-level social psychology theories of behavior and motivation, and macro-level political science theories of the causes of genocide. While it is a highly generalized model, Eris can be calibrated to match system-level conditions in different scenarios, contributing to the field by allowing for empirical validation across cases.

DEDICATION

This dissertation is dedicated to my mother, Mary Ellen Hastings Gust. Her unconditional love and support formed the bedrock of my life. She did not live to witness this achievement; however, she is with me every step of the way.

ACKNOWLEDGEMENTS

I extend my deepest thanks and appreciation to all members of my committee. It was absolutely one of the greatest honors of my life to work with this talented and diverse group of researchers who each showed an understanding and appreciation for the value of this work. Their time and expertise were invaluable, and the end result is not my dissertation, it is *our* dissertation. To begin, this research would not have been possible without the encouragement and support of my advisor, Mirsad Hadžikadić. Our mutual passion for the study of complexity and the prevention of genocide were the initial spark that began the process of this research. My co-advisor, Samira Shaikh, also provided essential guidance in bringing together different aspects of the research, particularly with respect to the issue of finding appropriate data for model validation. John Cox was our resident genocide historian and advocate for all oppressed peoples, ensuring that our understanding of history was sufficient to inform the work. Amy Canevello was our social psychology expert. She was not initially aware that her understanding of micro-level motivations would become an essential part of the model framework, and making this connection was one of my favorite moments in the research process. Min Shin contributed creative ideas and perspectives that improved the quality of the work. Finally, Bill Rand joined us as an external member. His support and feedback related to the model and validation process were perfect in every respect.

There were many members of the UNC Charlotte community who were critical to my success. From my colleagues in the Complex Systems Institute, to professors, administrators, support staff, and my students, each had a role in supporting me as I walked through the long process of earning this PhD. There are so many, I will not list names. You will all hold a special place in my heart as I move on to my next adventure.

I would also like to acknowledge the essential financial support I received from various organizations. UNC Charlotte provided me with GASP funding on a yearly basis, and awarded me the Joanna R. Baker Memorial Fellowship as well as the Graduate School Summer Fellowship. I received yearly external funding from the Fadel Educational Foundation, and

Foundation for the Carolinas. I am extremely thankful for all of the above assistance, as it would have been significantly more difficult to focus on my studies without this support.

My small circle of friends and family who have been with me through many joys and trials were always a source of strength. It has been an incredible journey, and I cannot imagine what would have become of me without all of you. Finally, I would like to thank each of my five children: Noura, Mohammed, Maryam, Iman, and YaSeen. You are the anchors of my life, giving me the ultimate reason to carry on through the most difficult of circumstances, and joining me in the celebrations of all things beautiful. We have experienced so much together, joy and pain, success and failure. I am honored to be your mother, to witness your evolution, and to have your company as I experience my own.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABM An acronym for Agent-Based Model.

ABMs Plural of ABM.

CAS an acronym for Complex Adaptive Systems.

CSS An acronym for Computational Social Science.

CHAPTER 1: INTRODUCTION

This dissertation introduces Eris, an agent-based model of identity-based conflict with the potential for genocide, named after the Greek goddess of "chaos, strife, and discord" [1]. The research is motivated by the ever present need to understand and prevent genocide. We begin by clarifying how this overriding motivation leads to a computational model, followed by related research questions, key findings, contributions of the work, and an overview of the full dissertation.

1.1 Motivation

The research path leading to the Eris model begins with complexity theory as it relates to understanding how micro-level behavior leads to emergent outcomes in human social systems. History and the social sciences provide essential theoretical understandings of human behavior in this context, and computational social science brings complexity and social science theory together into a model framework. The model's ability to reproduce general and specific dynamics, known as model validation, determines its usefulness as a tool for those seeking an artificial world in which to explore identity-based conflict.

Complexity theory as a theoretical framework. Complexity theory, specifically as it relates to complex adaptive systems, provides an optimal framework for understanding human social behavior. Our social systems are difficult to accurately predict due to their non-linear nature. People interact and adapt in changing circumstances, potentially bringing about more change, with that change continuing to spread to other people. These local interactions and adaptations can give rise to emergent phenomena such as stock market crashes, viral news, stampedes, genocides, and more. Complexity theory focuses on the nature of micro-level interactions within a system, seeking to carefully understand the relevant attributes and interaction rules that may lead to emergent outcomes. This theory forms the foundation of the research presented here, informing the exploration of social science theories for understanding the problem of genocide, as well as guiding the computational approach for applying these theories in a synthetic world.

The problem of genocide. Genocide causes unspeakable trauma, and the effects of that trauma are often felt by many future generations. Violence against people because of the identity with which they are associated continues today, as evidenced by the rise of hate crimes in the United States [2], genocide committed against Yazidis in Iraq [3], genocide committed against Myanmar's Rohingya population [4], and more. It is essential that research of past and present events continue in order to understand the causes of identity-based conflict, more accurately predict when and where it is likely to occur, and more effectively inform social and policy interventions such that they strengthen a given society's protections of all its people regardless of their chosen or imposed identities.

In "To Kill a People: Genocide in the Twentieth Century," John Cox defines genocide as follows:

> Genocide is the attempt to destroy any *recognized, stable, and permanent group* as it is defined by the perpetrator; it is a concerted effort to eliminate its individual members and to destroy the group's ability to maintain its social and cultural cohesion and, thus, its existence as a group. The perpetrators' genocidal intent can be uncovered by examining policies, actions, and outcomes [5, p. 11].

Many definitions are more specific as to how a "group" is defined, and what constitutes "destruction," including that of the United Nations [6]. However, given the abstract nature of a computational model, Cox's very general and inclusive definition is appropriate and we have selected it to guide this work.

Political science theories. Scott Straus theorizes that in addition to understanding known genocides, we also have a great deal to learn from negative cases. Positive cases are when genocide was expected and did occur, while negative cases are when genocide was expected and did *not* occur. A fundamental aspect of his theory revolves around the presence of "factors of escalation and restraint" in explaining genocide. The dynamics of restraint as understood by Straus heavily influence the approach of this research toward empirical model validation [7, 8].

In addition to yielding a model that achieves validation according to subject-matter experts by reproducing known behavioral patterns, another goal of this research is for the Eris model to be customizable such that it can reproduce known dynamics of different historic cases. Straus looks at a number of positive and negative cases from Sub-Saharan Africa, and we have elected analyze two of these: the 1994 Rwandan genocide as a positive case, and Côte d'Ivoire in the 2000s as a negative case. While the Rwandan example is well known and studied, Straus argues that the fact that Côte d'Ivoire did *not* devolve into a full-scale genocide in the 2000s is under-examined. [7, 8]. Reproducing the contrasting outcomes of these two cases is the goal of the Eris model's empirical validation process.

Social psychology theories. Complexity theory states that it is essential to consider what is happening on an individual level if we are to fully understand the emergence of events like genocide, which leads us from macro-level political science theory to social psychology theories of individual behavior within groups. Ervin Staub uses Social Identity Theory to explain conflict around the issue of identity, and he highlights the importance of bystanders in these scenarios [9]. At an even lower level, Crocker & Canevello's "Ecosystem and Egosystem Theory of Motivational Orientations" provides a framework for understanding how people are motivated by the fundamental need for self- and species-preservation [10]. The works of these researchers inform the lowest level of the Eris model, the point at which one person is interacting with another in circumstances that may lead to genocide. These interactions influence behavioral choices, and behavioral choices influence outcomes.

Computational social science. In the not so distant past, researchers examining identity-based conflict and genocide were restricted to the use of the historical record and its associated data in order to understand the past and predict what might happen in the future under similar circumstances. The accuracy of their predictions, and the effectiveness of any interventions, were only measurable after the fact. For example, those who saw the warning signs of the 1994 Rwandan genocide were only proven correct after the genocide occurred. In contrast, the warning signs identified in Côte d'Ivoire in the 2000s proved insufficient, and the expected genocide did not happen [7]. These researchers had few tools at their disposal that could aid their analysis and decision-making.

Computational social science is a research domain in which theories of human behavior are explored with computer models, and it is through this approach that we move from social science theory to the Eris model. Computer generated "artificial worlds" are a tool that can allow testing of hypotheses, as well as scenario and policy analysis, outside of the constraints of the historical record, and without the requirement of current or future human suffering. As computational power and resources have improved over the years, so has the ability to create powerful simulations of all kinds. In order to model discrete human interactions, Eris is an agent-based model (ABM). ABMs are one type of model used in computational social science. They are powerful tools for creating artificial worlds based on individual, interacting people. The theories that justify the model components and logic have explanatory power if they generate emergent phenomena in silico [11].

Having discussed the basic motivations and framework for the Eris model, we turn to specific, guiding research questions. These questions clarify the importance of choosing the correct type of computational model, selecting appropriate data that can be used for empirical validation, and determining the overall quality and usefulness of the work.

1.2 Research Questions

The broad goal of this research is to build a computational model of genocide, or more generally, of identity-based conflict. The following are three research questions that guide the process of reaching that goal. The first relates to the model itself, the second to data that can inform the model, and the third to how the model is evaluated and validated.

1.2.1 Research Question 1: A Computational Model of Genocide

What is the optimal, small and efficient, computational model of identity-based conflict, potentially leading to full-scale genocide, that will be useful in providing a means for unveiling and exploring the social dynamics of this problem and how these impact emergent outcomes?

Sub-questions:

- What is the optimal methodology for creating a computational simulation of identity-based conflict?

The goal of this research is to implement a small and efficient agent-based model that will be useful to researchers studying the problem of identity-based conflict that may lead to genocide. Historians and social scientists who conduct research in this domain must rely on analysis of real-world events in order to construct theories of why this type of conflict occurs and how it can be prevented. An artificial world that simulates human behavior in the context of identity-based conflict can provide an environment in which researchers can test their theories, obtain results that inform the next research cycle, compare those theories against others, and achieve all of this while avoiding the risk of human suffering. The importance of the work is clear; however, achieving the stated goal requires carefully determining the optimal methodology.

1.2.2 Research Question 2: Data For Empirical Validation

What are the most useful data to inform and empirically validate a generalized, agent-based model of identity-based conflict?

Sub-questions:

- In order to study the historical record, is it possible to extract text from data that will be useful for informing model parameters in order to achieve empirical validation on historic events.

The choice of data used to inform the model is a critically important issue, as the decision here will help determine the validity and generalizability of the research. This research will explore the use of historical data in the form of presidential speeches from both Rwanda and Côte d'Ivoire as they transitioned through instances of social turmoil, to genocide in the former, and to near-genocide in the latter [12]. These data will be analyzed using sentiment analysis to quantify elite emotion as it relates to motivation, thus providing a system-level measure of societal conditions that is related to the "Ecosystem and Egosystem Theory of Motivational Orientations" [10].

1.2.3 Research Question 3: Evaluation and Validation of the Implementation

Does this computational model provide an artificial world that is useful for social science researchers, policy makers, and others exploring the problem of identity-based conflict. Is it sufficiently small and efficient, and is it understandable and usable by those researchers?

Sub-questions:

- Is the validation meaningful?
- Can it allow for counterfactual analysis?
- Is it flexible between scenarios?

The expert opinion of historians and social scientists will ultimately determine whether or not the proposed model achieves its goals. In acknowledging that this research question is *theirs* to answer, we have worked to ensure historians and social scientists have a critical role in the development of the Eris model. Their expertise was, and is, essential in order to correctly interpret results, refine model parameters, and determine how to best analyze the data used for empirical validation.

1.3 Key Findings

The following are the main research findings for this work:

- The hypothesis that the ecosystem motivational orientation can be quantified from elite speech was affirmed by the results presented in Chapter 5.
- System-level factors of restraint as implemented in the Eris model have an exponentially beneficial effect on out-group death rates.
- Active bystanders do not need to be a majority in the environment in order to provide substantial protection to the out-group. With sufficiently high system-level factors of restraint, passive bystanders can remain a majority, and active bystanders will still be able to prevent out-group persecution.
- The Eris2 model, presented in Chapter 7, can be calibrated to reproduce general patterns of a real-world scenario. The model reproduced the approximate death rate of the 1994 Rwandan genocide.
- After determining the level of factors of restraint aligning with the Rwandan death rate, the Eris2 model was capable of modeling a relative adjustment of these factors based on the mean ecosystem findings for the Ivorian case presented in Section 5.3.3 and Table 5.5. The model showed that while the smallest gap in mean values was insufficient to prevent genocide, violence was greatly decreased for a median value, and nearly eliminated for the maximum possible difference in means between the two scenarios. The effect can be negated by increasing the perpetrator count to be approximately 25% of the out-group size.

1.4 Summary of Contributions

- The model moves social science research in the area from qualitative to quantitative in a novel way.
- The Eris model framework is an artificial world in which to explore the difficult problem of genocide. It achieves this by taking an approach informed by complexity theory and complex adaptive systems, which yields a simple model that provides useful quantitative measures.
- Eris is a generalized agent-based model of identity-based conflict and genocide that

captures commonalities across societies such that it can be easily calibrated for different real-world scenarios.

- The ecosystem and egosystem framework allow this agent-based model to operationalize the behavior of people. This novel connection between the fields of natural language processing and computational social science allows for empirical validation across cases.

1.5 Overview of Dissertation

This research seeks to make a positive contribution toward the study of identity-based conflict through the use of agent-based modeling. It is clear that the topic is of utmost importance given the grave consequences of the world's failure to prevent genocide. The model and methodology presented in this dissertation are the result of a close interdisciplinary collaboration, and it is the goal of all contributors to the work to ensure that the end results have value and positively impact the domain of conflict and genocide studies. The structure of the dissertation from here is as follows:

- **Chapter 2: "History and Social Science Literature Review"** Provides historical information regarding the two empirical validation cases: Rwanda and Côte d'Ivoire, describes the research methodology at a high-level, and reviews literature from social science theories that inform the Eris model framework.

- **Chapter 3: "Computational Modeling – A Theoretical Framework and Literature Review"** Reviews complexity theory, complex adaptive systems, model validation, and computational social science as foundations of the model's theoretical framework. Reviews computational models of conflict, identifying gaps in prior work this research addresses, and closes by combining what is presented in both Chapters 2 and 3 into a model framework.

- **Chapter 4: "Eris – A Model of Identity-Based Conflict"** Presents the components, logic, and verification of the Eris model. The chapter details the model environment, agent types and their characteristics, interaction rules, model limitations and key assumptions, results of the verification process, and introduces the experimental sets covered in the following two chapters.

- **Chapter 5: "Data"** Details the use of sentiment analysis to explore emotional content in the selected data. Next, provides an overview of the presidential speech data set, process of digitization, results of the sentiment analysis, and closes with a framework for transforming the results such that they can be used for empirical validation of the Eris model.

- **Chapter 6: " Experimental Set 1 – Generalized Inter-Group Dynamics"** Presents the methodology, results, and analysis of experiments run with the Eris model in a general format. These experiments explore inter-group interactions and emergent dynamics in a purely hypothetical society.

- **Chapter 7: "Experimental Set 2 – Reproducing Real-World Scenarios"** Reviews the results from the empirical validation process applied to the Rwandan and Ivorian cases. The work here combines the model as presented in Chapter 4 with the results of the data analysis from Chapter 5.

- **Chapter 8: "Conclusions"** Reviews and summarizes the research, key findings, and contributions. Presents a summary of limitations of the research, the future research agenda, and closing thoughts on the work as a whole.

CHAPTER 2: HISTORY AND SOCIAL SCIENCE LITERATURE REVIEW

In developing simulations of human systems, it is critical to have a solid understanding of the problem explored. To properly contextualize this work, we begin with a discussion of the history of genocide in Section 2.1. In the next sections, we take a closer look at the dynamics of genocide, both at the macro- and micro-levels. For the macro-level, political science theories are outlined in detail in Section 2.2, and micro-level theories from social psychology that explain individual actions and motivations are covered in Section 2.3.

2.1 History of Genocide and Identity-Based Conflict

The work of historians is the first step in the development of the Eris model. Historians do not require computational models to perform their work. It is the scientist seeking to understand and model the phenomenon of genocide who must rely on the records of historians, leaving him or her indebted to those who have worked to preserve and clarify the history of these horrific events. We begin with the definition of genocide and this is followed by a discussion of war, identity, perpetrators, and historic cases that inform the approach to data and model validation.

2.1.1 Defining Genocide

In order to properly frame any discussion of genocide, it is essential to provide its definition. Raphael Lemkin studied and contemplated the Armenian genocide and other episodes in which entire human groups were targeted for destruction. After fleeing Nazi-occupied Europe, and in light of the Nazis' crimes against Jews, he worked to clarify and develop a term for this specific type of violence and persecution. Lemkin recorded the crimes the Nazi party committed against the Jewish people and other marginalized groups, and coined the term "genocide," in "Axis Rule in Occupied Europe" [13]. Following his lead, the United Nations worked to clarify and codify what constitutes genocide. Many drafts, debates, and diplomatic bargains led to a definition without the more "universal, inclusive concepts" Lemkin believed should be included [5, pp. 4-6]. Article II of the 1948 "Convention on the Prevention and Punishment of the Crime of Genocide" provides the following definition:

> In the present Convention, genocide means 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 [6].

This definition was seen as flawed even by its authors, and scholarly debate on the issue continues today. Rather than providing a universal condemnation of persecution of *any* identity group, the Convention's definition lists only four categories, which leaves political, sexual, and other groups unrecognized and unprotected [5, p. 6]. To address the shortcomings of the UN definition, many have worked to develop alternatives. John Cox's definition, quoted in Section 1.1, is the formal definition selected for this research; however, there are many others to consider. Adam Jones provides a chronological list of many such definitions in "Genocide: A Comprehensive Introduction" [14]. Jones' list begins in 1959 with Peter Drost: "Genocide is the deliberate destruction of physical life of individual human beings by reason of their membership of any human collectivity as such," and ends in 2009 with Donald Bloxam: "[Genocide is] the physical destruction of a large portion of a group in a limited or unlimited territory with the intention of destroying that group's collective existence" [14, p. 16-20]. These two examples show that genocide can be understood in very general terms, providing recognition of *all* social groups.

Computational models are, by their definition, abstractions of reality. As such, this work does not attempt to contribute to the discussion of what constitutes genocide and what does not. Rather, the model constructed here aims to be as general as possible. The goal is to build a computational representation of any type of conflict that occurs around the issue of identity. It is up to the user to determine if the scenario applies to the problem of genocide.

2.1.2 War, Identity, and Perpetrators

The connection between war and identity is fundamental to genocide, and should therefore be explicitly addressed. Merriam-Webster provides several definitions of war, including: "a state of usually open and declared armed hostile conflict between states or nations" and "a struggle or competition between opposing forces or for a particular end" [15]. The former specifies that the conflict is "armed" in nature, and the latter is more inclusive in terms of the acting groups, expanding beyond states and nations.

War, or armed conflict of some kind, is a common factor in nearly all cases of genocide [5, pp. 204-205][16, p. 194][17]. In clarifying what constitutes genocide, it is important to determine how conflicts that lead to this outcome differ from other types of war. Wars are fought for a wide variety of reasons, including territorial expansion, access to natural resources, or an internal quest for power or regime change. Genocide becomes a possibility during wartime when identity, or more specifically, *social* identity, is a salient factor. Social identity is determined by an individual's association with a group, and genocide occurs when there is an attempt to destroy that group. As noted in Cox's definition of genocide presented in Chapter 1, it is only necessary for the *perpetrator* to apply the identity to the victim, marking that individual as a target for persecution [5, p. 11].

Perpetrators are the agents of violence and are typically members of sub-national groups, which are distinct from the elite power structure as discussed in Section 2.2 below. Martin Shaw defines these as "armed power organizations." He clarifies that these groups can range from the highly organized and well armed, to the unarmed with the ability to "mobilize" those who are able to commit violence [16, pp. 194-195]. Going further, the 1994 Rwandan genocide showed that even unusually large groups of civilian actors can take part in genocidal violence [18, p. 95]. Finally, although most genocides are the work of an identity group with a numerical majority, it is also possible for "subaltern" groups to commit these types of atrocities [19]. While the nature of the actors may vary considerably by situation, it is the intent and actions of the perpetrators that advance genocide.

The elimination of an identity group is often not the initial motivation for a particular war; however, genocide can occur when a group is perceived as a threat or used as a scapegoat [9]. These variables and relationships are discussed in detail in Sections 2.2 and 2.3.1 below. History tells us that there are two critical factors that can make genocide a possibility: war of some kind, and polarization of societies around the issue of identity. Positive cases have these, and the circumstances have led to horrifying genocidal violence. Negative cases also have these escalating factors to some extent; however, other factors in the societies appear provide restraint against large-scale violence and genocide. This differentiation between case types is informed by Scott Straus' work, which highlights the importance of examining both, as cases in which an expected genocide did not happen may also contain valuable information for helping prevent future violence [7, 8]. We next review two historic scenarios: the positive case of Rwanda in 1994 , and the negative case of Côte d'Ivoire in the 2000s. These selected events serve to explain the problem and its variation, as this is critical to a well informed computational model. For the reader who wishes to delve more deeply into the history genocide, there are many excellent books on the subject [5, 14, 20, 21, 22, 23, 24, 25, 26].

2.1.3 Positive Case – Rwanda (April–July, 1994)

Despite efforts to ensure that the global community would not allow the development of genocidal violence, it has occurred again in more than one location. The 1990s saw genocide both in Southeastern Europe and Sub-Saharan Africa. In this section, we examine the case of Rwanda in 1994, which was distinguished by its horrific daily death rate and unusually high level of civilian participation in the killing [14, pp. 346-362][27].

Rwanda is a poor, landlocked country located in East Central Africa. Its neighbors are:

Democratic Republic of the Congo to the west, Uganda to the north, Tanzania to the east, and Burundi to the south. Rwanda was colonized by Germany in 1894, and then Belgium by 1916 when Germany was no longer able to retain control. Prior to the colonization of their country, Tutsi and Hutu were terms that described "social status," with the former associated with a higher status of "pastoralists," and the latter considered a lesser class of "cultivators." While this differentiation of class status enabled the Tutsi to be a more powerful group despite the fact that they were a minority of the population, there was no biological determinant for group membership. In fact, acquiring cattle was an avenue for a Hutu to become a Tutsi [5, pp. 153-154].

The Germans, and more so the Belgians, developed a polarizing framework for Rwandan society. They "racialized" the differences between the Tutsi and Hutu, elevating the Tutsi to a privileged position despite their status as a minority in the overall population. In the European tradition of transforming ethnic and class divisions to *racial* divisions, the colonial powers imposed a system of identity cards, segregation, and other methods to create rigid boundaries between groups where they had previously been fluid [5, p. 154]. Jones states that the "divide-and-rule tradition" of colonialists was typical, and motivated the creation of this polarized ethnic structure [14, pp. 348-350].[1]

The privilege and dominance bestowed upon the Tutsi caused "resentment and vengefulness" amongst the Hutu population. After World War II, the Belgians chose to begin elevating Hutu over Tutsi. The late 1950s and early 1960s were a period fraught with tensions and violence between ethnic groups as Rwanda moved toward independence. Rwanda's first president, Grégoire Kayibanda, imposed extreme levels of discrimination against Tutsis [18, p. 23]. This led to "proto-genocidal massacres" of thousands of Tutsis, and the forced exile of tens of thousands of others. President Juvénal Habyarimana came to power by coup in 1973, and while his anti-Tutsi policies were less extreme than those of Kayibanda, he ruled Rwanda in dictatorial fashion, closely consolidating power and continuing to sow ethnic di-

[1] Note that Rwanda has a third ethnic group. The Twa ethnic group represents approximately 1-2% of the total population [14, p. 350],[28]

visions. He "...gradually increased ethnic hatred against the Tutsis, encouraging a climate of fear and panic to forestall demands for democracy" [14, p. 350].

Meanwhile, Tutsi exiles were organizing and arming themselves across the border in Uganda, ultimately forming their own resistance army in 1987, the Rwandan Patriotic Front (RPF). In 1990, the RPF invaded Rwanda from the north, leading to several years of increased tensions, armed conflict, and killing of Tutsis. This ultimately led to the signing of the Arusha Peace Accords in August of 1993, and the introduction of 2,500 UN peacekeepers from the United Nations Assistance Mission for Rwanda (UNAMIR). However, "Hutu Power" extremists within Rwanda could not abide these developments, and escalated their efforts to prepare for a potential massacre of all Tutsis [14, pp. 350-351].

President Habyarimana's plane was shot down on April 6, 1994, and this was the spark that lit the genocide. Estimates of the number of Tutsi and moderate Hutus killed in the following 13 weeks vary. Some figures are as low as 500,000, and others reach almost 1,000,000, with Tutsis represented more than 90% of the victims overall [14, p. 360]. The world did not act to protect the people of Rwanda even at the height of the killing.

General Roméo Dallaire was the commander of the UNAMIR peacekeeping force, and has struggled to live with the fact that he and his small force of peacekeepers were unable to prevent this genocide. He recorded his experience in "Shake Hands with the Devil: The Failure of Humannity in Rwanda." He states "...I was unable to persuade the international community that this tiny, poor, overpopulated country and its people were worth saving from the horror of genocide–even when the measures needed for success were relatively small," and that "...the Rwandan story is the story of the failure of humanity to heed a call for help from an endangered people" [29, pp. 515-516].

It is clear that even in the 21st century, "Never Again" is a goal not yet reached. However, all hope is certainly not lost. In the next section, we examine a case in which genocide could certainly have happened and did not.

2.1.4 Negative Case – Côte d'Ivoire (2000s)

In "Making and Unmaking Nations: War, Leadership, and Genocide in Modern Africa," Scott Straus brings together his years of research comparing and contrasting positive and negative cases of genocide [8]. He finds negative cases are underexamined, and that they may contain critical information about factors that make a society less likely to experience genocide. One of these negative cases is that of Côte d'Ivoire in the 2000s.

Côte d'Ivoire is located in West Africa. The Atlantic Ocean is to its south, Liberia and Guinea to its west, Mali and Burkina Faso to its north, and Ghana to its east. It was colonized by the French, and Félix Houphouët-Boigney was its first post-colonial president from 1960 until his death in 1993 [30]. As president, Houphouët-Boigney strove to increase the nation's economic growth, which meant cultivating the sparsely populated and very fertile land of the west and southwest. Continuing existing French policies, he strongly encouraged internal migration and immigration from other West African countries. These policies and others led Côte d'Ivoire to become economically strong, with its capital Abidjan becoming known as the "Paris of West Africa." Throughout his presidency, Houphouët-Boigney "...emphasized the importance of tolerance, solidarity, prosperity, and discourse." In addition to his support of ethnic diversity and the stability provided by a long-term, strong economy, Straus finds that the nature of this "founding narrative" was an essential restraint to group-selective violence and genocide in later years [8, p. 123-151].

Despite its stability and success, Côte d'Ivoire was not immune to economic and political crises. In the late 70s and through the 80s, global commodity prices dropped, which strained the country's economy due to its reliance on exports of cocoa and coffee. In the background were informal land use arrangements that made the status of land ownership unclear, causing tensions between people native to certain areas, and Ivorian and West African immigrants. In the political realm there was a global push for democracy, leading to the introduction of additional political parties in Ivorian elections starting in 1990. These and other factors, coupled with the death of Houphouët-Boigney in 1993, pushed Côte d'Ivoire into crises that brought about two civil wars in the 2000s [8, p. 127-131].

Henri Konan Bédié became president after Houphouët-Boigney's death until he was overthrown in 1999 in a coup led by General Robert Guei. This was followed by years of struggle for power between political parties, leading to the involvement of the French and international peacekeepers [30]. Through these crises, a "nationalist ideology" emerged. Ivoirité elevated the Akan ethnic group after Houphouët-Boigney's death. This ideology "crystallized anti-foreigner, anti-northerner, and anti-Muslim sentiment in the southern and western parts of the country." Straus finds that this ideology also served as a justification for "nativist" claims to fertile lands in areas populated by other identity groups. Localized violence against marginalized ethnic groups during two civil wars led international organizations to warn that genocide was highly likely in the early 2000s and later, after the 2010 elections [8, pp. 123-125][31].

What prevented genocide in Côte d'Ivoire? Straus cites a number of restraining factors including the strong intervention of the French, West African nations, and UN peacekeepers. However, despite the fact that many experts state that these interventions were the most critical element of deescalation, Straus convincingly argues that there were many additional restraints integral to Ivorian society. The most significant of these was the lack of widespread and systematic dehumanization of any particular identity group. The wars were not seen as conflicts between identity groups, and there was little "ethnic hatred." Ivoirité was largely seen as "...a cultural idea used by politicians to win elections, to win over the people at the base. It was not in the minds of Ivorians." In other negative cases of genocide, Straus notes commonalities with respect to the larger narratives of the elite. Emphasis on "pluralism, multiethnicity, and dialogue" are, in his view, defining features of a society that can avoid genocide even in the most difficult of circumstances [8, pp. 138-168].

2.1.5 From History to Social Science

The above history of genocide, its dynamics, and examination of select cases, are an effort to convey relevant information sufficient to contextualize the computational work. Before proceeding, we present Figure 2.1, which visualizes our high-level methodological approach.

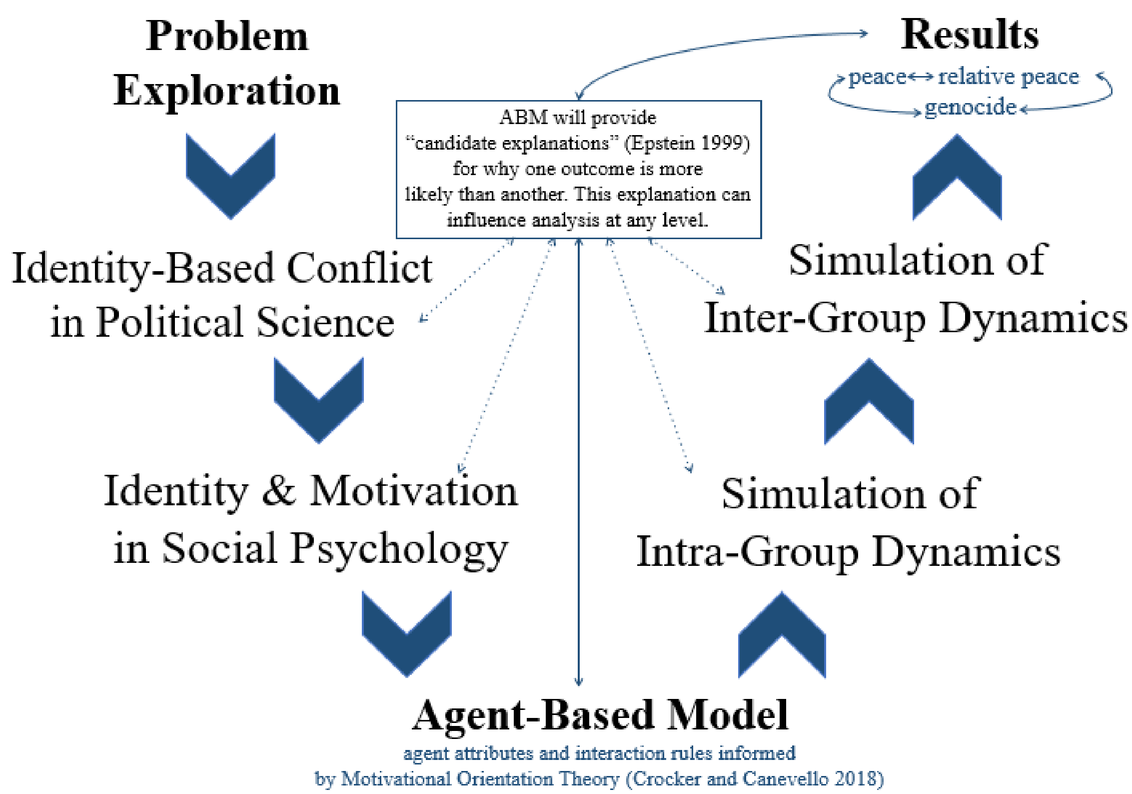


Figure 2.1: High-level visualization of general research methodology.

As shown in the diagram, we begin with problem exploration. This process moves from the system-level understanding of identity-based conflict as informed by political science, to research specific to identity and motivation as understood by social psychologists, to the implementation of these theories in an agent-based model. The computational model is theoretically grounded through this process, yielding a small and efficient model with theory-based attributes and parameters. Simulations of intra- and then inter-group dynamics have the potential to yield results that show the evolution of peace, genocide, or states in-between, with the theories that inform the ABM providing potential explanations for what causes or restrains genocide, influencing analysis at any level.

We next examine political science theories that explain the system-level dynamics of genocide. These theories help inform the model environment as described in Section 4.2.

2.2 Political Science Theories

Political science is the study of states, their institutions, and their governance. Research in this area generally uses systematic and empirical methods. Present day approaches are interdisciplinary, relying on theories from fields like psychology and sociology in order to better understand and explain political phenomena [32].

In "Modeling Genocide at the System and Agent Levels," von Briesen, et al. visualize the connections between different elements of a system at risk for genocide as shown in Figure 2.2 [33]. This simplified causal loop diagram primarily relies on the work of political scientists Scott Straus [7, 17] and Ernesto Verdeja [34, 35], and psychologist James Waller [36].

In a causal loop diagram, variables of interest are connected to each other by arrows that indicate the polarity of their relationship. A '+' sign at the end of an arrow signifies that the two variables have a positive causal relationship, meaning that when one increases, the other does as well. A '-' sign signifies that the two variables have a negative causal relationship, meaning that when one increases, the other decreases. Positively related variables move in the same direction, and negatively related variables move in opposite directions [37, p. 139].

Elite-Civilian Relationship. The diagram in Figure 2.2 includes several feedback loops of interest. First, 'R1' is the feedback loop between the level of ideological extremism among a society's elite and how compliant its in-group civilian population will be. The letter 'R' indicates that this is a reinforcing loop, as both variables have a positive causal relationship. This means that as the societal elite become more extreme in their ideology, civilians become more compliant, allowing the elite to become more extreme given there are no restraints on the feedback mechanism. This relationship between elite actors and civilians is the diagram's central loop, highlighting its importance to the overall process of developing genocides. Scholars vary in their interpretation of the degree of impact this relationship between elite and civilian sectors has on a genocidal outcome; however, it is generally recognized to be an important factor [17, 34, 36].

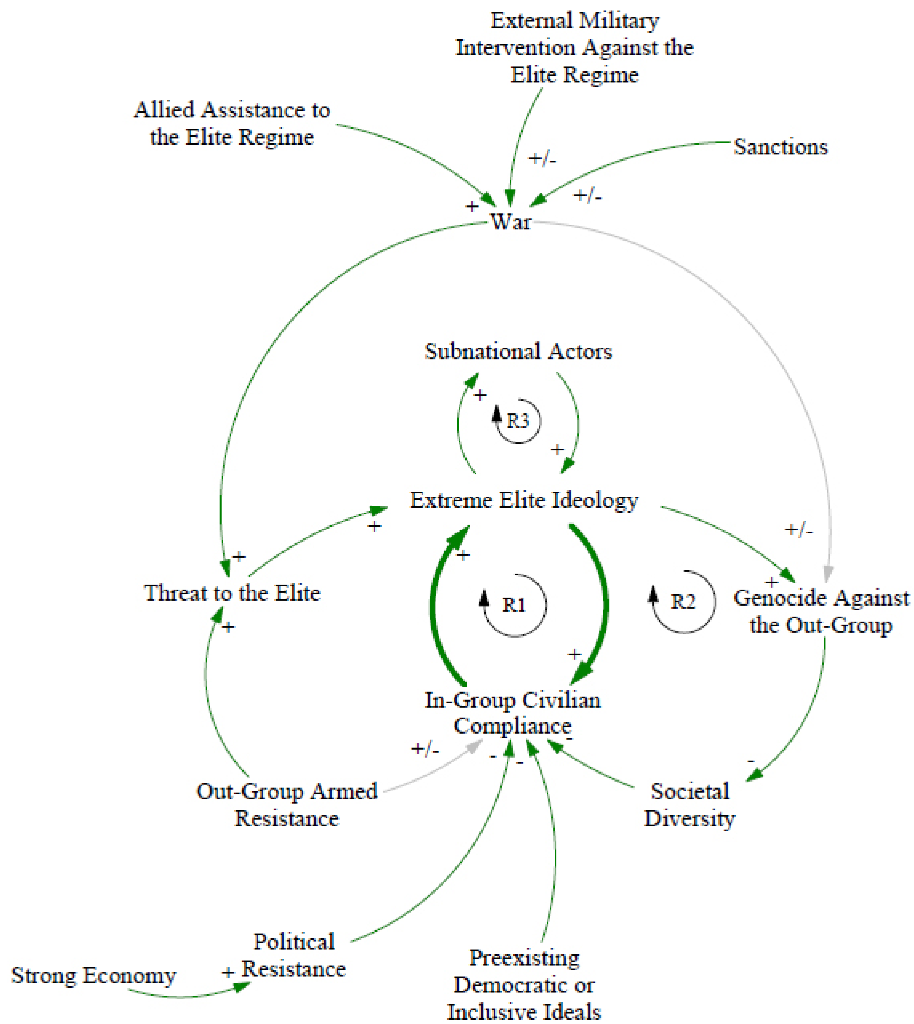


Figure 2.2: Visualization of system-level dynamics of genocide [33].

The cycle of genocide. 'R2' indicates a larger reinforcing feedback loop connecting variables that drive genocide. As the elite becomes more extreme, they can plan for and execute genocide against an out-group. Cox states that genocide is:

> ...enduringly destructive: The annihilation, partly or wholly, of a group diminishes the diversity and richness of the human race. When a people is eradicated or dispersed and its traditions and culture erased, all of human civilization loses much that can never be regained [5, p. 189].

This effect is factored into the feedback loop by showing that as society becomes less diverse due to genocidal violence, fearful civilians are more likely to be compliant, allowing the elite to become more extreme. While this research has not yet fully validated this relationship, it is included due to the likelihood of its importance.

The elite-subnational relationship. Straus finds that the majority of genocides are committed by subnational actors rather than civilians [17]. Rwanda was an exception in this regard, with a high level of civilian involvement in the killing [18, p. 95]. In Figure 2.2, the supportive relationship between subnational actors and the elite forms reinforcing loop 'R3.' The more extreme the elite ideology becomes, the more support they receive from their subnational counterparts, thus allowing extremism to build. Examples of subnational actors are local governmental officials, military leaders, and so on. This research takes a simplified understanding of these relationships, recognizing that there may be cases in which subnational actors are not loyal to the regime. Examining the complex relationship between the elite and subnational is left for future research.

Factors of restraint. Connected to the "In-Group Civilian Compliance" variable in Figure 2.2 are several factors that can act as restraints against increasing compliance. In his examination of negative cases of genocide, Straus pays particular attention to societal conditions he theorizes have provided restraints against increased violence and persecution of out-groups, and the diagram focuses on two particular restraints from his work. First is a "Strong Economy" that leads to "Political Resistance." Straus finds that negative cases of genocide often occur in societies with a strong middle class, and theorizes that their influence on the elite power structure acts as a restraint, since with armed conflict comes great uncertainty and instability. "Preexisting Democratic or Inclusive Ideals" is another restraining factor he highlights in his research. The established presence of these types of ideals, in long-standing democracies for instance, increases the likelihood that in-group civilians will resist developing extremism [7, 8].

Factors of escalation. Straus finds that "war and forms of exclusionary ideology" are the most important factors of escalation that can drive genocide [7]. War, along with out-group armed resistance, are incorporated into the diagram as variables that increase the level of threat to the elite regime, thus contributing to a rise in extremism. Allied assistance to the elite regime also has a positive relationship, as it has a strong potential to increase the scope of war and success of an extreme elite regime. As discussed in Section 2.1.2, the presence of war and armed conflict in general are common across all cases of genocide. As such, these factors of escalation have a strong, positive causal relationship with an extreme ideology that can motivate group-selective violence.

Factors with uncertain effects. Every scenario is different, and there are a number of variables with uncertain effects, as evidenced by the historical record. Allied intervention in World War II caused the Nazis to escalate their genocidal efforts; however, it turned the tide of the war against the Axis powers, eventually ending the genocide [5, p. 109]. Sanctions can also have a complex effect. While the intent of economic and political sanctions is to intervene without engaging in warfare, these can have unintended consequences. Jones state that the "...difficulty with sanctions lies in targeting them to impede a repressive or genocidal leadership, without inflicting general human suffering" [14, p. 573]. If sanctions can effectively restrain war and extremism, they can help to prevent genocide. However, it is possible that they will cause deprivation and suffering of the in-group, leading to a higher potential for out-group persecution.

The emergence of genocide. Genocide does not occur in a vacuum. As explained above, Straus finds that societies form different types of relationships between in- and out-groups, and other factors can ease or strain those relationships. Returning to the positive case of Rwanda covered in Section 2.1.3, we can see that the plane crash killing President Habyarimana was what Straus and Verdeja term a "trigger" of genocide [38, 39]. The presence of civil war as a factor of escalation. Minimal restraints against out-group persecution in the society, coupled with this triggering event, led to the emergence of one of the worst cases of genocide in human history. However, it is important to recognize the role of restraining factors in preventing genocide. Triggering events during the Ivorian civil wars did *not* lead to genocide, leading Straus to theorize that understanding the dynamics of restraint is essential in understanding how to prevent this type of violence [7, 8].

Connections to the computational model. The theories of political scientists are crucial in developing a realistic and valid model of genocide. They provide a high-level framework for an artificial world with a potential for genocide. In the case of this research, factors of restraint play a critical role in the model. As explained in Section 4.2, the model allows the user to create an abstract representation of these through the use of a system-level function.

While the above provides essential understandings and frameworks, societies are composed of human beings. Underneath the layers of government are individual people interacting with one another. As discussed in Section 3.1, complexity theory tells us that emergent phenomena are the result of micro-level dynamics. Therefore, this research turns its focus to in-group civilians, and how their compliance or non-compliance affect an extreme elite agenda (see feedback loops R1 and R2 in Figure 2.2). If the elite require a compliant civilian population in order to further a genocidal plan, what determines whether or not these civilians cooperate? We next explore the social psychology of individuals in the face of extreme circumstances. Why do they act as they do? Why do many become fearful? Why do some resist? What have psychologists found are salient factors that can help answer these questions?

2.3 Social Psychology Theories

As explained in Chapter 1, this research develops an agent-based model (ABM) of genocide. An ABM is a model of *individuals*; therefore, it requires an understanding of individual traits and interactions. These traits are realized at the micro-level of a society, and influence the reactions and responses of individuals to what is happening in their environment. In short, genocide is the result of the social interaction of human beings.

One must understand the psychology of the individual in relation to society in order to properly capture the underlying dynamics of identity-based conflict. Social psychology is a field that studies behavior as a function of the self and the environment [40, p. 30]. It is the "scientific study of the behaviour of individuals in their social and cultural setting" [41]. Theories here attempt to answer questions about the underlying causes of individual actions and reactions within their society or social setting. Next, we explore questions about the nature and effects of evil and extraordinary situations. These are followed by the examination of two particular theories that inform the computational model: Social Identity Theory, and the Ecosystem and Egosystem Theory of Motivational Orientations.

A great deal has been written about the role of individual personality in the presence of extreme violence. In "Becoming Evil: How Ordinary People Commit Genocide and Mass Killing," James Waller thoroughly explores crucial questions regarding how individuals allow extraordinary violence and "evil" to occur [42]. Many have sought to identify a definitive personality type that causes an individual to perpetrate evil. We can rest more easily if we believe that we, as individuals, are not capable of committing genocidal acts. Psychologists administered Rorschach tests to imprisoned Nazis seeking to find evidence of such a personality type or related traits, and did not succeed in identifying any specific "mental disorder" common to all subjects [43]. Christopher Browning writes in detail about the "ordinary" personalities of a German police battalion in "Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland." He finds that "ideological indoctrination" and "conformity to the group," among other factors, were sufficient to understand the violence these men committed against Polish Jews [44, pp. 184-186]. While it may disturb us to learn that extreme circumstances may cause any of us to act in ways we would normally consider unimaginable, understanding our human weaknesses can help us more effectively work to prevent future genocides [45].

Gustave Le Bon's seminal work, "The Crowd: A Study of the Popular Mind," details his observations of social behaviour during the French Revolution. He states: "In a crowd every sentiment and act is contagious, and contagious to such a degree that an individual readily sacrifices his personal interest to the collective interest" [46, p. 7]. Human social systems are complex, as discussed in Sections 3.1 and 3.3. Feedback mechanisms that arise during interaction can cause authority figures to influence small groups of people, who form small crowds, which then influence neighboring groups of people, and so on. The two theories explained next help distill these complex interactions such that they can inform a computational framework by providing the discrete elements of individual identity and motivation.

2.3.1 Social Identity Theory

While individuals may have a personal understanding of their own identity, such as being male or female, a morning person or a night person, and so on, they also have identities defined by their relationship to others. Social identity theory works to understand how individuals identify with groups, and how that impacts their perceptions of people from other identity groups [47]. Ervin Staub clarifies the difference between "personal" and "social" identity by stating that personal identity is connected to how the individual answers the question "who am I?" In contrast, social identity is defined by "...the extent to which individual identity is rooted in, or connected to, the group." He goes on to clarify that "...in difficult times people try to strengthen their identity through personal identification with some group" [9]. Related to Le Bon's observations of the behavior of crowds is identity fusion theory, which helps explain how individual identity can merge with the identity and actions of the group. This theory states that when people experience a strong connection to their group, they "...channel their personal agency into group behavior, raising the possibility that the personal and social self will combine synergistically to motivate pro-group behavior" [48]. Given this high-level understanding of social identity, we next cover additional aspects of identity and its influence on processes leading to genocide.

The dimensionality of identity. While the computational model developed in this research simplifies the representation of identity, it is in fact complex and often multidimensional [49, 50]. Most people do not consider themselves, and are not considered by others, to be part of only one identity group. As an example, a Mexican American born in the US may see himself as Mexican *and* American, affiliated with a particular religion, part of a particular social class, and more. Each of these represents a dimension of that person's identity. However, how he is seen by others can be relevant as well, particularly when a group is being marginalized by other, more powerful groups. Even if he does not see being of Mexican heritage as being relevant, others may apply that identity *to* him, particularly if they wish to target that group for derision or persecution.

The salience of identity. Given that identity can be multi-dimensional, it is important to understand that different dimensions can have different levels of salience. For example, a child born to two Irish parents may be considered "Irish" by her family, even when she decides that ethnicity is not personally relevant. Here, being identified as Irish is salient to the family, and not to the individual. In the case of genocide, this issue takes on a heightened importance. It did not matter to the Germans if a Jew did not see himself as affiliated with that group. Neither did it make any difference if a child was born of a Tutsi mother and Hutu father if she was identified as Tutsi by her killers [5, p. 154]. In short, social identities are complex and held at many levels, all the way from the individual to the global community.

The dynamic processes of genocide. As opposed to other types of violence, genocide is characterized by focused attacks on a particular identity group within a society. Figure 2.3 below is an adaptation of Ervin Staub's visualization of genocide dynamics [9] included in von Briesen et al.'s recent work [51]. This diagram shows the generalized, escalating process that is what Staub calls a "continuum of destruction" [9][52, pp. 17-18]. Instigating events like war, economic crisis, and more, cause real or perceived threats to the basic needs of a people. In seeking to obtain safety and security in uncertain and difficult times, groups may strengthen their internal bonds, which can increase polarization in some societies. Harmful actions toward marginalized groups can then escalate, leading to societal changes sometimes so dramatic that the process leads to genocide. Staub clarifies that this process is characterized according to the culture of the society in question. For example, a lack of electricity would threaten the "basic needs" of most residents of the United States; however, this is a common condition in areas of Sub-Saharan Africa [53]. Finally, Staub highlights the importance of bystanders in the process. Their willingness to "allow" the escalation violence and destruction is essential for the evolution of genocide [9]. As such, this research elects to closely examine their role and the influences on their behaviors in order to inform the Eris model.

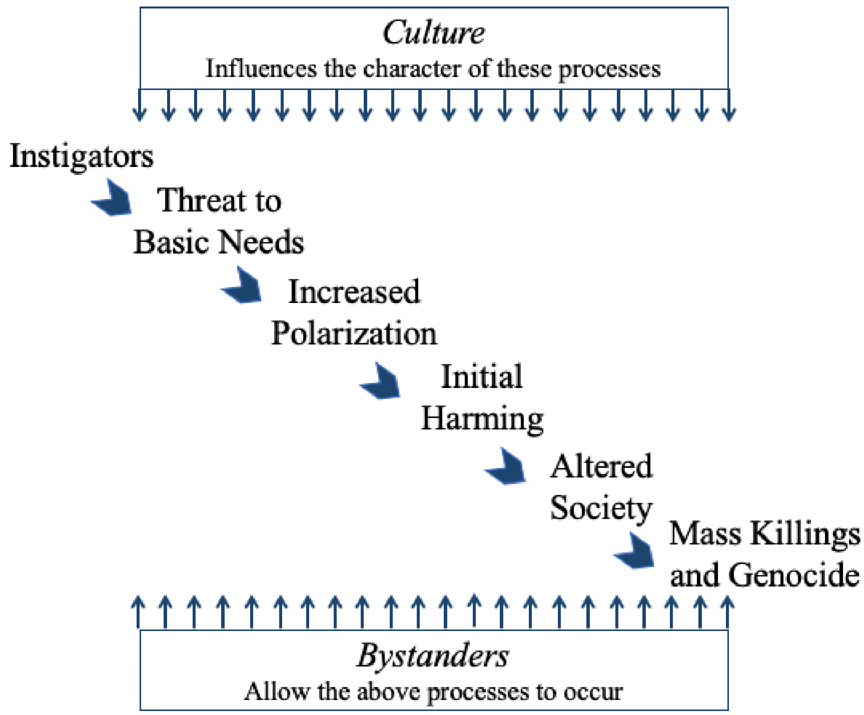


Figure 2.3: Visualization of the evolution of genocide in [33], adapted from [9]

Bystanders in identity-based conflict. Returning to Figure 2.2, we see that a compliant in-group civilian population is necessary in order for feedback mechanisms to reinforce one another, allowing a progression toward genocide. Staub places bystanders in three categories: passive, active, and complicit. Table 2.1 outlines how Staub defines each of these categories [54, pp. 13-36].

Table 2.1: Descriptions of bystander types as defined by Staub [54]. (Table first appears in [51].)

Bystander Category	**Description**
Active	Can be found "...speaking out in behalf of their values and in behalf of people who are harmed" (pg. 33).
Passive	Stand in contrast to active; however, do not engage with perpetrators in a way that makes them complicit.
Complicit	Do not intentionally "...support harmdoing but by their actions, make perpetrators believe that at the very least they accept what they do" (pg. 14).

The existence of a compliant civilian population implies that its bystanders are either passive or complicit. Active bystanders are resisting in some way, advocating for those who may be in harms way. Staub argues that actions of active bystanders can influence the outcome, and that their behaviors can, in some instances do, determine a genocidal or non-genocidal outcome [52, 54]. The leads to the question of how are people motivated to intervene in behalf of the oppressed and victimized. Why would a person who considers him or herself from identity group A put forth any effort to stand in solidarity with a person from identity group B? Moreover, why would he or she do so in the presence of conditions leading toward violence against group B? The next section explores a motivational theory that provides a framework for the lowest level of agent motivation in the Eris model.

2.3.2 Motivational Orientation Theory

Figure 2.4 visualizes Crocker & Canevello's "Ecosystem and Egosystem Theory of Motivational Orientations" [10]. This theory provides a framework for understanding how individual motivations relate to our "evolved motivations" for self- and species-preservation. Crocker & Canevello find that egosystem motivation is characterized by "self image goals," and ecosystem motivation by "compassionate goals." These systems operate independently of one another, influencing the nature of the support individuals give to others.

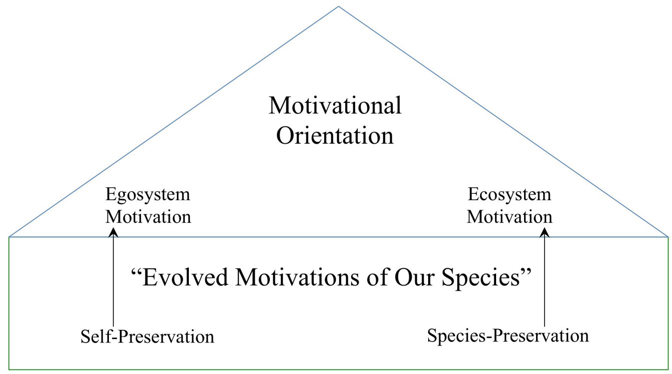


Figure 2.4: Scaffolding of motivations according to The Ecosystem and Egosystem Theory of Motivational Orientations [10]. (Figure first appears in [51].)

When oriented toward the egosystem, Crocker & Canevello find that the individual has "self-image" goals. Here, people are motivated "...to be seen by others as having desirable characteristics, and not be seen as having undesirable characteristics." This leads to affective states of competitiveness, conflict, confusion, and fear. In contrast, ecosystem orientation is characterized by "compassionate" goals that bring about "...caring for the well-being of others and energizes behaviors to protect them and support them to thrive." Crocker & Canevello's research finds that individuals with ecosystem motivation have less anxiety, and tend toward feelings of cooperativeness, peace, love, and clarity [10].

2.3.3 Theoretical Framework for Connecting Social Psychology Theories

Returning to Staub's theory of bystander behavior during identity-based conflict, we now have a point of connection between his characterization of active bystanders and individual motivations. Staub finds that active bystanders are altruistic and prosocial [54, pp. 35-36]. These traits are what one would expect given an ecosystem orientation, as a person oriented in this direction is motivated to help and support others in a way that is compassionate and *not* directed toward the self. Table 2.2 details a framework that formally connects social identity theory, as it relates to bystander behavior, and motivational orientation. The hypothesis here is that the goals and affective states found in egosystem and ecosystem motivational orientations influence the probabilities of individual bystander decisions.

Table 2.2: Framework for connection individual motivational states with probabilities of choice of bystander types [51].

Motivational Orientation	**Goals**	**Affect**	**Bystander Classification**
Egosystem	Self-image	Competitiveness Conflict Confusion Fear	High Egosystem $\vdash$ (yields) Higher probability of Complicit Bystander
Ecosystem	Compassionate	Cooperativeness Peace Love Clarity	High Ecosystem $\vdash$ (yields) Higher probability of Active Bystander

We have arrived at the end of the problem exploration portion of the research methodology shown in Figure 2.1. The next step is to transform the above theories of human behavior into the abstract world of a computational model. Computational approaches can aid research in this area by producing calculations, data analysis, predictions, as well as simulated realities for problem exploration. The framework outlined in Table 2.2 is the ground-floor of what will become Eris, an agent-based model of identity-based conflict and genocide. Chapter 3 begins with an exploration of complexity theory, clarifying its relevance to this work (Section 3.1). This is followed by a description of the characteristics and uses of agent-based models to study complex systems (Section 3.2). Section 3.3 presents an overview of the field of computational social science, and then explains the use of agent-based models to simulate social systems as a way of contributing to social science research. The chapter closes with a review of prior computational models of conflict (Section 3.4), identifying relevant research gaps, leading to a novel model framework to explore the problem of identity-based conflict 3.5.

CHAPTER 3: COMPUTATIONAL MODELING – A THEORETICAL FRAMEWORK AND LITERATURE REVIEW

The previous chapter reviewed the problem of genocide through the lenses of history and the social sciences. The next step is to connect these theories to a computational framework. Section 3.1 explores complexity theory as an anchor for the application of computational tools to the problem of genocide as qualitatively defined in the social science literature. Section 3.2 discusses the use of agent-based models for the study of complex systems, and is followed by an overview of the field of computational social science in Section 3.3, particularly theory driven work in this area using ABMs to study social systems. Section 3.4 reviews a variety of models of conflict, identifying research gaps and areas for contribution, and Section 3.5 brings all of the above together into a theoretical framework that informs the model presented in Chapter 4.

3.1 Complexity Theory

Researchers across disciplines have long been interested in understanding human social phenomena. The motivations behind their work vary according to their personalities, backgrounds, and areas of expertise; however, they all encounter similar difficulties in understanding human systems due to the seemingly infinite diversity and complexity of humans themselves. Computational approaches to modeling human behavior and systems are continually advancing. When model development is coupled with a theoretical understanding of the non-linear nature of these systems and their human residents, the models can successfully expose previously hidden layers and dynamics. Complexity theory has a strong potential to yield useful explanations of what is often unpredictable human behavior. This potential stems from its focus on micro-level attributes and interactions; when these are present in a system of agents, they often lead to emergent phenomena. In the case of humans, emergent "social" phenomena can range from war, to traffic jams, to stock market crashes, and beyond.

3.1.1 Foundations of Complexity Theory

It appears the human impulse to understand and explain the natural world motivated Alan Turing to develop his 1952 theory of morphogenesis [55]. Philip Ball summarizes Turing's inspired application of his skills by stating that he "...devised a mathematical model that explained how random fluctuations can drive the emergence of pattern and structure from initial uniformity" [56]. Bringing his mathematical genius to the domains of biology and chemistry, Turing states that his goal was to "...discuss a possible mechanism by which the genes of a zygote may determine the anatomical structure of the resulting organism." One of his underlying questions asked: how can systems with "spherical symmetry" evolve and produce organisms, such as horses, that are not "spherically symmetrical" [55]? Ball explains Turing's model as showing how: "...chemicals that are diffusing and reacting may produce neither bland uniformity nor disorderly chaos but something in between: a pattern." While this was not a unique concept at the time, Ball claims that Turning was the first to "...relate such an abstruse phenomenon to the question of biological growth and form-in short, to suggest how chemistry alone might initiate the process that leads from a ball of cells to a starfish, a horse or to us" [56].

Turing's work is one of earliest modern contributions to the development of complexity theory. In this research area, complex systems are understood to be non-linear, with macro-level properties that cannot be explained by a simple aggregation of the system's micro-level components [57]. These macro-level manifestations are frequently termed "emergent phenomena," and are of particular interest to those studying complex systems. As Turing hypothesized, some global constructs cannot be explained simply through analysis of the micro-level structure of the given system. This logic not only applies to biological systems, but also extends to physical, social, and other system types. The common theme for all is that the whole is "greater than the sum of its parts," and this drives rich, fascinating, and often interdisciplinary complexity research.

3.1.2 Complex Adaptive Systems

To further refine this discussion of complexity theory, we turn our focus to complex "adaptive" systems (CAS). Micro-level learning and adaptation are the defining characteristics of complex adaptive systems. These systems have "...many distributed, interacting parts, with little or nothing in the way of a central control" [57]. This is a fine point, best explained through contrasting examples. For instance, a hurricane can be classified as a complex system due to its non-linear nature and composition of interacting air and water molecules. The emergence of this category of storm cannot be explained through reduction to its constituent parts, and notably, these micro-level components do not themselves learn and adapt. In contrast, human social systems can be examined from the perspective of individual people who are inherently able to adapt and learn as they interact within their environment. These interactions may lead to the emergence of tribes, nation-states, wars, peace, and so on. Holland states that "...standard theories in physics, economics, and elsewhere, are of little help because they concentrate on optimal end-points, whereas complex adaptive systems 'never get there.' They continue to evolve, and they steadily exhibit new forms of emergent behavior." These systems are fascinating to study, and the advent of modern computing capabilities has made their modeling and simulation increasingly realistic [57].

3.1.3 Computing in the Study of Complexity

Today's "massively parallel" computers can model and simulate systems with "many rules that are active simultaneously" [57]. One can only imagine where Turing might have gone with his inherent understanding of complexity were he given sufficient computing power to model interaction. Holland states that "[i]n seeking to adapt to changing circumstance, the parts can be thought of as developing rules that anticipate the consequences of certain responses." Even if the anticipated event never occurs, adaptation can bring about significant global changes. Price increases in anticipation of shortages that never occur are one example.

The "emergent ability to anticipate" is a critical dynamic "...that makes the emergent behavior of complex adaptive systems intricate and difficult to understand." These systems can now be computationally modeled using techniques such as system dynamics, agent-based, and mathematical modeling. While explanations of emergent phenomena can remain elusive, the computing "laboratory" is ever expanding in its capacity, flexibility, and power [57]. Next, we turn our focus to human social systems and how theories of complex adaptive systems can be explored through the use of computation to yield insights into emergent social phenomena.

3.2 Agent-Based Models (ABMs)

An agent-based model (ABM) has the unique ability to encode the micro-level interactions that are the foundational characteristic of the complex adaptive systems discussed in Section 3.1.2. ABMs do this by providing a computational framework in which the user can define the appropriate environment, its agents, and their respective attributes. In this artificial world, the encoded interaction rules between agents as they exist within their environment can lead to both expected and *unexpected*, emergent outcomes.

The critical difference between ABMs and top-down simulation techniques, like system dynamics [58], is that ABMs are modeled from the "agent perspective" [59]. While agents live within a global environment that may have conditions beyond the control of the individual agent, it is *individual* characteristics, behavior, and adaptation that motivate the model. Macal argues that this modeling approach stems from a desire to capture the "real world" as we intuitively understand it, from our own, individual perspective. Doing so yields simulations that provide insight into micro-level "causal mechanisms" that can be useful to those trying to understand the underlying causes of emergent phenomena in a system [59].

3.2.1 The Components of ABMs

What is an ABM? At their most basic level, ABMs are models composed of a heterogeneous set of autonomous agents with defined attributes, interacting inside a simulated environment according to local rules. These agents typically learn or adapt as they interact with one another, which can result in emergent macroscopic regularities of interest to the researcher.

Macal and North [60] provide a succinct and useful "tutorial" of ABM techniques, detailing the structures and composition of agents and their environments from which the researcher may choose. Agents may have static attributes such as race or gender, and/or dynamic attributes like wealth, memory and age. Examples of environment types are: "non-spatial," meaning that agents do not have defined locations, a fixed "grid or lattice," as in the case of cellular automata, a geographically realistic spatial representation, or a network in which agents are represented by nodes [60].

In his paper "Everything you need to know about agent-based modelling and simulation," Macal [59] outlines suggested definitions of ABMs. In each case, agents have a "diverse set of characteristics," with their interactions and potential adaptation being the elements that differentiate one type of ABM from another. His definitions are presented in Table 3.1, and help inform the following sections explaining each of the fundamental elements of an ABM: the environment, agents, and interaction rules.

Table 3.1: Definitions and descriptions of types of ABMs. Adapted from [59].

ABM Type	**Level at which agent behaviour is determined**	**Interactions**	**Adaptation**
Individual ABMs	Exogenous: system-level	Limited	None
Autonomous ABMs	Endogenous: agent-level	Limited	None
Interactive ABMs	Endogenous: agent-level	Between agents and each other. Between agents and the environment.	None
Adaptive ABMs	Endogenous: agent-level	Between agents and each other. Between agents and the environment.	Adaptation is possible.

What is the environment and what are its attributes? At its most basic level, an ABM environment is the world in which its agents reside. Its attributes are system-level characteristics, and are defined according to the problem being studied. For example, an ABM of stock market dynamics could simulate a trading floor, where a system-level attribute of importance could be management directives, or global economic conditions. Alternatively, an ABM of neighborhood gentrification could have a very detailed environment that represents a real-world city, including system-level attributes such as local zoning ordinances, access to public transit, and more.

What are agents and what are their attributes? As stated above, agents in an ABM are autonomous. They have individual attributes which may be homogeneous or heterogeneous as compared to other agents in the simulation. An agent is the element that differentiates an ABM from other types of models, as it is defined at the micro-level. This allows the modeler fine control of the lowest levels of a system, allowing him or her to create artificial societies composed of appropriate synthetic agents. Examples of agents are cars in traffic [61, 62], birds in flight [63], people in a city where a contagious disease is spreading [64], or immune and non-immune cells as they allow for tumor growth [65].

What are interaction rules? As outlined in Table 3.1, agent behaviour in an ABM is determined endogenously in all cases except for an Individual ABMs, where this is determined at the system-level. Individual and Autonomous ABMs have limited agent interaction, while Interactive and Adaptive ABMs allow agents to interact with each other and/or their environment. In Interactive and Adaptive ABMs, rules that define what happens as agents interact with one another and/or their environment are codified abstractions of the scenario in question. Of the possible types of interaction in these models, we focus on two: agent-agent and agent-environment [66, pp. 257-262]. Agent-agent interaction can range from simple predator-prey interaction [67], to the act of sensing the state of other agents in a flocking model where one bird detects the distance and heading of another [63], to the act of communication in a social network [68]. Agent-environment interaction causes changes to the state of the environment. Returning to Schelling's segregation model, agent actions cause the environment to change such that segregation develops due to preference-based agent movement [69]. In a modified version of this model, agents could account for environmental policies in making their decision of where to move next. An ABM that allows for the types of interactions described above encodes rules that specify different behaviors and results as agents move through the simulated world.

What is adaptation? Adaptive ABMs from Table 3.1 are the only class of model that allow for temporal changes to agent state and behavior [59]. There are many ways in which to incorporate these types of changes in an ABM. Some models implement the agent ability to have memory and learn, while others encode an increased probability of survival for higher levels of agent fitness for the environment, and some models combine these two or implement other methods of allowing states and behaviors to change in time. As an example, the "Traffic Basic Adaptive" model allows the user to explore both adaptive and non-adaptive behavior. In the non-adaptive mode, a car determines if it should speed up or slow down based on the behavior of nearby cars. When allowing for adaptation, this car is able to determine the most optimal "acceleration value" from the past, and use that information to determine their best speed within the current environment [70]. Given the importance of adaptation in many complex systems, it is often the case that an adaptive ABM is the best choice for abstracting the system in question.

What can results can an ABM give? ABMs provide artificial worlds in which one can explore systems built from the bottom-up. Given carefully designed micro-level components, appropriate interactions, and adaptation if it applies, a researcher can use the model to explore these systems through a series of experiments. Later in this chapter we review examples of agent-based models and their results. Section 3.3.4 provides and overview of seminal ABMs of social systems, and Section 3.4.2 reviews ABMs specific to conflict.

How are ABMs implemented? For agent-based model implementation, NetLogo [71] is the development tool of choice for many researchers. It allows for rapid prototyping, and provides a useful graphical interface that aids model interpretation. Macal and North list other options for constructing ABMs, including development environments like MASON and Repast, and general object-oriented programming languages such as Java, C++, and Python [60]. More recent options include Mesa, a Python library for building agent-based models [72], and NL4Py which allows the user to control a NetLogo workspace through a Python client [73]. It is important to note that while these tools are essential, the foundational choices of granularity, structure, and elemental composition will determine the final results, their validity, and their usefulness.

3.2.2 Model Validation

What is validation and why is it important? In the field of modeling and simulation, validation is a process used at many levels to determine how well a given model corresponds to reality. This process determines if a model is an accurate and meaningful representation of a known system. Rand & Rust provide a concise outline of the different types and purposes of validation, and the sections below are based on their work [74].

"Face" validation. Face validation means that the model's components correspond to the real-world on the micro- and/or macro-level. It is critical that the modeler ask the appropriate questions to ensure that the model is, on its face, realistic [74]. As a micro-level example, one can ask if the rules for individual birds in flight are informed by valid research in avian behaviour. At the macro-level, one can question how well the system and its global attributes and rules reflect a real-world trading floor in an ABM of stock market behaviour. Much of the process of determining face-validity at any level depends upon the input of subject-matter experts. For example, Zhai et al.'s model of online collective emotions in a crisis relies on organizational theories to inform micro-level attributes and behaviours [75]. Subject matter experts in the domain of organizational communication and behavior are essential in determining whether or not their theories have been properly translated into this computational model.

Empirical input validation. Many models of real-world systems use empirical data to inform their initial parameters and behaviors, and it is important to ensure that these data accurately correspond to the problem in question. Methodologies for validating data used to inform ABM model parameters are continually developing. Zhang & Vorobeychik's recent review of input validation processes for ABMs of innovation diffusion provides an excellent example of the complexity and variation of this evolving process. The authors find that different modeling approaches use different types of input data, yielding a number of issues, some of which can be addressed by taking advantage of advances in data analytics [76]. A robust approach in all areas of agent-based modeling is essential when a model's results are intended to inform policy and other decision making. The first step toward a valid representation of a specific, real-world problem is ensuring that the data used to inform the model are both accurate and appropriate [74].

Empirical output validation. The basic premise of this level of validation is to ensure that the model's output corresponds to the real-world. Figure 3.1 visualizes the spectrum of output validation according to how each methodology connects the model to its goals. This diagram is adapted from von Briesen, et al. [51], which itself is informed by the work of Rand & Rust [74].

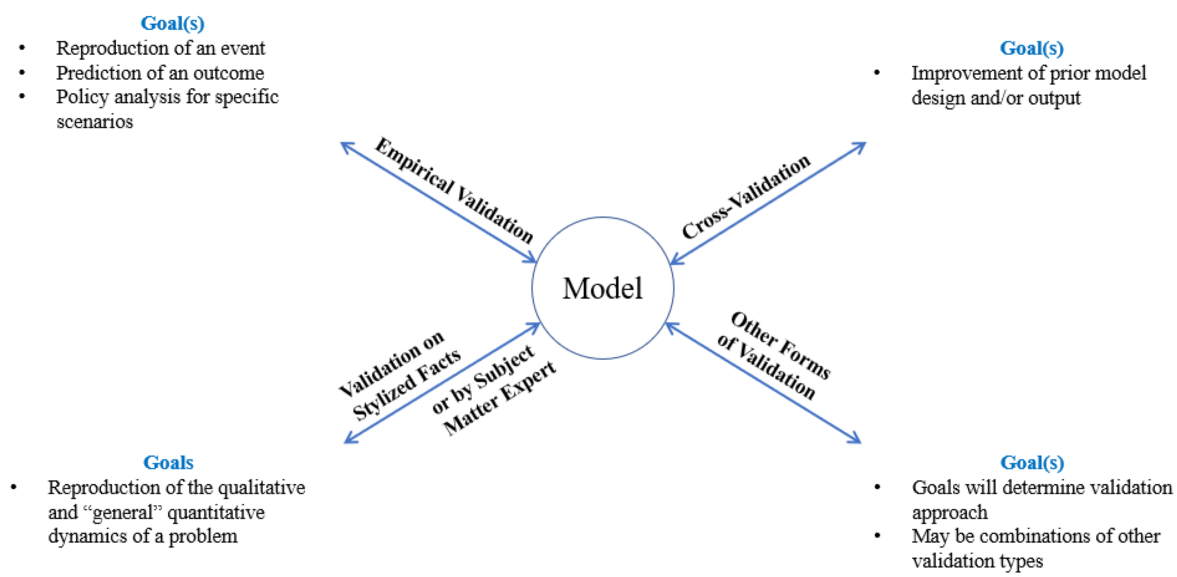


Figure 3.1: Forms of ABM output validation as they relate to the goals of the model in question. (Based on the work of Rand & Rust [74], diagram adapted from von Briesen et al. [51].)

As shown in the figure, there are four major output validation methodologies. An ABM with a goal of reproducing a real-word event, or predicting an outcome, requires empirical validation. That model must be able to simulate known events according to event-specific data. In contrast, an ABM whose goal is to reproduce very general patterns and dynamics can be validated on stylized facts according to subject matter experts. These experts provide the constraints and measures necessary to ensure that the model correctly reproduces known dynamics. When the goal of the model is to improve upon a previously developed and validated ABM, cross-validation is the appropriate methodology. If the new model can produce results similar to those of the original, then it has a stronger validity. Finally, it is possible that the modeling goals are combinations of the above and will require a custom methodology. The model presented in this dissertation is an example of this type last approach, as its goal is to reproduce general dynamics of genocide while also being able yield scenario specific results [74].

3.2.3 ABMs of Complex Adaptive Systems

We conclude this section with a brief discussion of how and why agent-based models are used to explore complex adaptive systems (CAS) as described in Section 3.1.2. Recall that a complex adaptive system is a dynamic system composed of homogeneous or heterogeneous agents. These agents interact within their environment, with these interactions governed by rules set at the local/agent level. There can be multiple feedback mechanisms and opportunities for agents to learn and adapt as they interact with one another. Agents are also affected by system level conditions. The dynamics of these elements often leads to self-organization of agents and emergent system properties such as non-linearity, power-law dynamics, chaotic cycles, and others. Emergent properties are a common characteristic of complex adaptive systems, and their defining quality is that they arise from agent-level interactions and are not imposed from the system level downward. CAS are studied using a variety of modeling techniques, including those found in the fields of mathematics, system dynamics, and network science. Agent-based modeling is appropriate when a dominant feature of the system is its governance by local, agent level, interactions. While many of these systems can be well understood and modeled using mathematics, ABM techniques introduce an additional layer of examination both by fine-tuning agent rules, and by allowing for qualitative measures that cannot be represented using other techniques. Using ABMs to explore human social systems as complex and adaptive leads us to the field of theory-driven computational social science. Research in this domain has led to exciting new ways to encode low-level human interaction, adding to our understanding of the reasons behind emergent outcomes.

3.3 Computational Social Science

From Schelling's model of segregation [69] to Epstein and Axtell's "artificial society" known as Sugarscape [77], the "computational study of social phenomena" [78] has been a continually evolving area of research. When the work is an interdisciplinary collaboration between social scientists and computational experts, it brings the strengths of both together in order to explore the complexities of human social systems. Research in the domain of computational social science (CSS) is of utmost importance, because as society grows and changes, so do its problems. Gilbert & Bullock highlight the potential of this work with respect to the connection between the physical and social sciences, including computer science. They state that "...research at the 'social sciences interface' has the potential to transform our ability to answer questions about important social, socioeconomic, socioecological, and sociotechnological systems" [79].

In the field of theory-driven computational social science (CSS), theories of micro-level interaction guide the development of computational models intended to simulate human social systems. This section begins with an overview of the field, then discusses how the CSS research has policy making significance, and closes by defining and explaining the importance and benefits of producing social models that "generate" macro-level phenomena from micro-level foundations that are clearly defined and grounded in social science research [11].

3.3.1 Overview and Definitions

Computational social science is an emerging field that is generally comprised of two areas: data-driven [80] , and theory-driven [81]. Data-driven CSS uses computational tools to explore social realities as encoded in data artifacts. Predictive algorithms can use that data for training, and then forecast potential outcomes. In the area of conflict research, this commonly manifests as risk scores. These algorithms and associated data quantify how similar a current scenario is to one from the past, and thus how probable it is that there is a risk of violence. Section 3.4.1 reviews these models in detail. Alternatively, theory-driven CSS typically involves the use of simulations informed by established theories of human behavior and interaction. If the simulations are able to reproduce known behaviors, the theories behind the models become what Epstein terms "candidate explanations" for emergent outcomes [11]. Data are often used in this work, as discussed above in Section 3.2.2; however, social science theory is the fundamental driver of model development.

3.3.2 The Computational Approach of Generative Social Science

The study of social systems will naturally "...involve dealing effectively with large-scale complex systems made up of many parts interacting and adapting in sometimes subtle ways over significant timescales" [79]. Joshua Epstein, a seminal researcher in the field of CSS, states that "[m]any important social processes are not neatly decomposable into separate subprocesses–economic, demographic, cultural, spatial–whose isolated analysis can be somehow 'aggregated' to yield an adequate analysis of the process as a whole" [11]. Epstein is describing complex adaptive systems, as discussed in Section 3.1.2, and this is one point at which computational modeling and complexity theory intersect. Mathematical models are often insufficient to capture the dynamics of such systems in a way that sheds light on the underlying interaction mechanisms. Computational modeling techniques, such as agent-based modeling (ABM) discussed in Section 3.2, allow researchers to build simulations of social systems from the bottom up, creating more interpretable models that are useful for understanding micro-level dynamics and how they lead to global, emergent phenomena. With these understandings, CSS researchers have made, and continue to make, significant contributions to society by aiding in the understanding of the origins of large-scale social phenomena.

Epstein & Axtell propose that computational social science should evolve in a "generative" direction:

> Artificial society modeling allows us to 'grow' social structures in silico demonstrating that certain sets of microspecifications are *sufficient to generate* the macrophenomena of interest. And that, after all, is the central aim. As social scientists, we are presented with 'already emerged' collective phenomena, and we seek microrules that can generate them. We can, of course, use statistics to test the match between the true, observed, structures and the ones we grow. But the ability to grow them–greatly facilitated by modern object-oriented programming–is what is new. Indeed, it holds out the prospect of a new, *generative*, kind of social science [77].

Building computational models with micro-level attributes and mechanisms that then "grow" macro-level dynamics creates new possibilities for discovering and analyzing those very microspecifications. Epstein postulates that "...if the microspecification m does not generate the macrostructure x, then m is not a candidate explanation. If m does generate x, it is a candidate" [11]. Determining these "candidate explanations" is the main purpose of generative social science, as computational models that yield these become laboratories for exploring human social systems.

More recently, Conte and Giardini [78] express concern for the current state of affairs of CSS, as the "analysis of large amounts of data to find correlations among them, and possibly predict future courses of events" has overshadowed the importance of understanding underlying "behavioral mechanisms that generate" these correlations. They propose refocusing the trajectory of the field toward the "generative" approach described above, combining "social data mining and computational modeling" in order to more comprehensively understand and analyze systems. As an example, Barbara Harff's statistical model of genocide discussed below in Section 3.4.1 could have a generative extension [82]. This would involve incorporating her findings into a full simulation of human society that accounts for interaction between people and/or nation states. Here, an agent-based simulation of genocide might require the use of Harff's results to realistically parameterize the model settings, while its agent-based nature would allow for incorporation of human behavioral and cognitive qualities relevant to understanding the micro-level dynamics of importance. If the resulting simulation "grows" a genocide, then the model's microspecifications may provide new and helpful insights into understanding, predicting, and preventing the phenomenon. A number of models have been developed along these lines, employing a variety of goals and structures. The next section reviews the basic modeling methodology in this domain, as well as two seminal models from the field of CSS.

3.3.3 ABMs in Computational Social Science

Bruch and Atwell review a number of agent-based models (ABMs) to stimulate discussion of the required "dimensions of realism" [83]. While their suggestions and conclusions are directed toward empirical social science research, they are important to consider within the full range of CSS modeling. They state that to "goal" of these models is "...not to reproduce empirical patterns or incorporate all aspects of reality so much as to understand the implications for social dynamics of one or more empirical observations or stylized facts." This is a critical consideration, as too fine grained an approach to CSS can result in model output that cannot be interpreted. They correctly point out that "[a] key issue for the analyst is the appropriate level of model complexity and empirical realism" [83]. The very nature of social systems is that they are complex and adaptive, as explained above in Section 3.1. Murray Gell-Mann, one of the founders of the Santa Fe Institute, a leading center for interdisciplinary research of complex adaptive systems [84], highlights the importance of choosing the right level of granularity by stating that "...when defining complexity it is always necessary to specify a level of detail up to which the system is described, with finer details being ignored. Physicists call that 'coarse graining' " [85, p. 29]. This is the art of modeling complex adaptive systems, and once established, the researcher can proceed to model development on solid ground.

3.3.4 Seminal ABMs of Social Systems

Schelling's model of segregation [69] is an early, micro-level model of this social phenomenon. At the time, his model could not be "computational" in its implementation, but it certainly had computational "logic" at its core. It is often cited as a seminal piece of research in CSS. Schelling showed that segregation could evolve from micro-level movement motivated by seemingly benign discriminatory preferences. While he did not have the ability to use a digital computer, his "...abstract study of the interactive dynamics of discriminatory individual choices" employed the type of computational logic we now implement in an agent-based model. In his most basic model, Schelling simulated people from different "groups" as stars and zeros, and placed them randomly along a line. Agents would "move" if their "neighborhood" was not composed of agents of their own type by at least half. His manual process of iteration revealed that the simple desire to avoid being in the minority was enough to produce distinct patterns of segregation. He also explored a wide variety of environmental conditions and agent preferences in both linear and two-dimensional models. The main takeaway from Schelling's work with respect to CSS is that he was able to create an abstraction of a real-world social system, extracting the essential components and interaction rules of importance, and he found that these were sufficient to produce segregation. His results highlight the effects of apparently innocuous, and some would say reasonable, individual preferences to avoid living as a minority in a neighborhood. This has provided inspiration not only for addressing the issue of segregation, but also for applying computational logic to complex human social problems whose nature and causes remains difficult to understand and determine.

Epstein and Axtell [86] produced one of the foundational works in the field of "generative" social science. They called their "artificial society" Sugarscape. At its most basic level, this model is designed to study the "distribution of wealth" among agents gathering sugar, a "renewable resource" distributed in the environment. A two-dimensional, 50x50 grid in the form of a torus, is the "sugarscape." Sugar is distributed on the grid such that there are two "peaks" in the northeast and southwest quadrants, with sugar capacity highest on the peaks and then dropping in gradient fashion, resulting in some areas that are void of sugar. As an environmental rule, sugar has a "growback rate" that determines how quickly it regenerates once consumed. Sugar can range from regenerating instantly, to growing back according to a complex set of locational rules [86, pp. 21-23].

Agents are populated randomly across the grid, and they are endowed with a "sugar metabolism" that defines how rapidly they deplete their gathered stock of sugar. Next, they have a "level of vision" that constrains their ability to assess and move through their environment. Finally, agents have "wealth," quantified in sugar reserves. Agents chose to move from their current location to a new site according to rules based on the maximum quantity of sugar available in their radius of sight. They collect available sugar when they arrive at a new location, and their sugar level is depleted according to their metabolic rate. If an agent cannot gather enough sugar to make up for the amount it consumes, it dies and is removed from the simulation [86, pp. 23-26].

Besides sugar, what else does this base Sugarscape model "grow?" Epstein and Axtell found that if sugar regenerates immediately to its full capacity when consumed, agents form populations along "ridges" in the "terraced" sugarscape. They deduced that these patterns form because agents' level of vision forces them into local optima provided they are able to find locations with sugar capacity that meets their metabolic needs. When sugar does not grow back instantly, and instead regenerates by "1 unit per time period," agents form "colonies" in the high sugar density areas of the grid. Epstein and Axtell were able to study the "carrying capacity" of the sugarscape, and showed that "selection" favored "agents with low metabolism and high vision" [86, pp. 26-32].

In their book length account of their development of and research with Sugarscape, Epstein and Axtell go on to describe how they refined environmental and agent attributes and interaction rules in order to model historical processes, trade, and disease transmission [86, pp. 54-152]. A full review of their findings is beyond the scope of this document; however, we take one as a representative example.

Epstein and Axtell simulated the transmission of "culture" as an element of their model of historical processes. To accomplish this, they endowed agents with cultural "attributes" simulated by a string of zeros and ones. Each position along the string is a "tag," and an agent's neighbors change their value (zero or one) of a randomly selected tag if its value does not match that of the original agent's corresponding tag. Given the encoding of culture and this "transformation rule," they then formed two "groups" whose members were determined by "majority vote," with agents belonging to the Blue group if zeros "outnumber" ones, and the Red group otherwise. Epstein and Axtell were interested in discovering if their artificial world would produce "spatially segregated, culturally distinct groups," and it did. With initial group membership and tag-flipping rules in place, the population on the sugarscape always converges to one group or the other. There may have been "wild fluctuations...en route to equilibrium," but that single group equilibrium was the consistent outcome nonetheless. The inclusion of sexual reproduction and combat allowed Sugarscape to "grow a crude caricature of early social history," indicating that historical processes may emerge from a very basic set of micro-level components [86, pp. 54-93].

These are representative examples of seminal, generative models from the domain of computational social science. Each model reveals the lengths to which its developer or developers have gone to distill the factors of importance in order to effectively explore the human phenomenon in question. Schelling left us with essential insights into the effects of seemingly "reasonable" preferences for levels of neighborhood diversity. Epstein and Axtell created a rich "artificial world" that shows how even something as complex as "history" can be generated from a few simple environmental and agent attributes and interaction rules [86].

3.3.5 Contributions of Computational Social Science to the Social Science Research Cycle

Concepts from complexity theory can be applied in the domain of computational social science in order to correctly determine relevant microspecifications, and then grow accurate and useful simulations of human systems. Seminal models are the first point of reference in order to to verify that this approach can lead to increased understanding of social phenomena. Next, we take a closer look at the contribution of computational models to the social science empirical research cycle as a whole.

Bhavnani et al. make an important contribution to the field of computational social science in “Rumor Dynamics in Ethnic Violence” [87]. The model and results detailed in the paper are very important; however, their explanation of the basic process of using ABMs to enhance empirical research provides a great deal of clarity in determining appropriate areas of focus, understanding limitations, and maximizing the power of these models to provide useful results. Figure 3.2 visualizes the basic process outlined by Bhavnani et al. This begins with social scientists performing empirical research, using their theories and results to inform a computational model, which then yields its own results. If the computational results confirm the empirical research, the model can be used to explore the dynamics of that problem, as it has proven to correctly simulate established “mechanisms and dynamics.” If the results do not confirm the empirical research they can be used to explore and inform options for future empirical research [87].

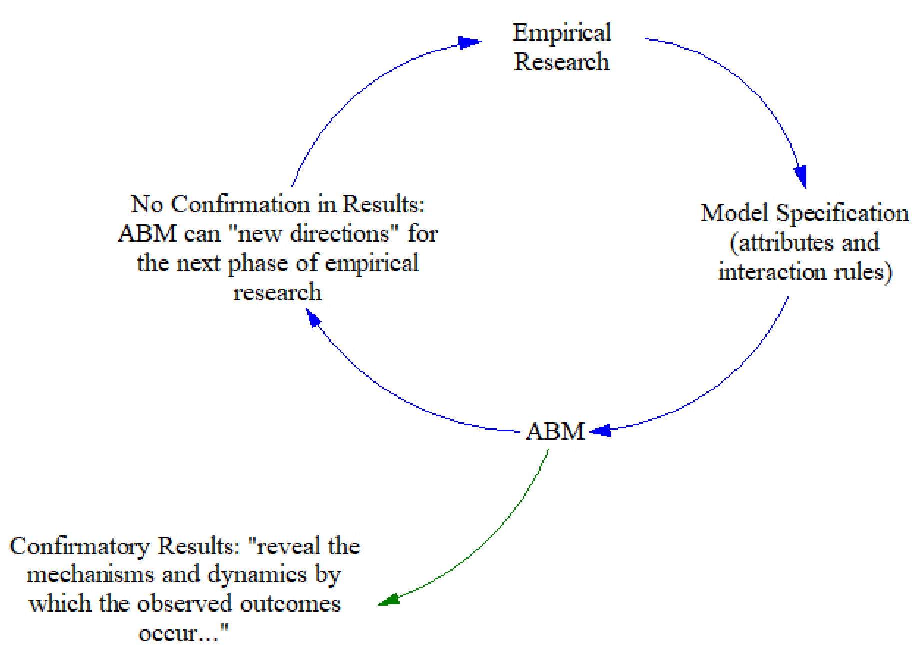


Figure 3.2: ABM contribution to empirical research cycle (based on the work of [87])

The above explains the contribution of a computational model to the empirical social science research cycle in a most general sense. A properly constructed ABM can provide social science researchers with an artificial world in which to explore and test theories based on prior empirical studies, or theories for which there exists no empirical data. This can also be done on a larger scale and over longer periods of time than may be possible in human studies. While this contribution does not remove the difficulties encountered in studying large human systems directly, such a model can help researchers test hypotheses and scenarios, refine research questions, and gain a better understanding of the system and its dynamics. All of the above can benefit the formulation of plans for the next human study. Before moving from a theoretical to an implemented model framework to study genocide, we next review relevant computational models of conflict and genocide.

3.4 Computational Models of Conflict – Literature

The models included here fall into two general categories: machine learning, and agent-based models. Machine learning models generally have a predictive goal, using data-driven techniques to provide output like that can predict the likelihood of genocide, or measuring the quantitative significance of different factors influencing a genocidal outcome. In contrast, agent-based models may or may not use data to drive the research questions and model framework. Many ABMs reviewed here are intended to explore general dynamics of conflict, while others are scenario-specific and intended for policy analysis in the given society.

3.4.1 Predictive Models of Conflict

Predictive models apply data-driven, statistical analyses to identify risk factors and forecast the likelihood of genocide and mass violence. They are designed to "learn" from data, and can then be applied to other scenarios in order to quantify the likelihood that the new case will yield results similar to those from the past. In "Predicting Genocide and Mass Atrocities," Ernesto Verdeja draws on the work of Barbara Harff to clarify two general categories of models in this area: risk assessment and early warning models [39, 88]. Both types of models yield predictions; however, the goals of those predictions are different. Risk assessment models examine "long-term structural conditions" as they relate to past atrocities, and therefore the risk of similar events in the future. Alternatively, early warning models seek "short- and midterm predictions" of violence [39]. The overall results of models in these categories are important because they provide quantitative measures that are often difficult for the human to infer. These measures can then be used as a supplement to the advice of experts in determining the variables of significance in conflict and genocide, as well as providing information regarding the risk of civilian violence in a particular society.

Risk Assessment Models. There are a number of major research projects and associated models in this area, and many of these rely on Barbara Harff's seminal work "No Lessons Learned from the Holocaust? Assessing Risks of Genocide and Political Mass Murder since 1955" [82]. Harff used data from 126 instances of "internal war and regime collapse" between 1955 and 1997 to create her model. Of these 126 instances, 35 were "geno-politicides," and she was able to isolate six significant factors that differentiated wars resulting in genocide from those that did not. These factors included political upheaval, prior genocides, and low economic development. Harff's analysis is extremely thorough, and has been widely cited by other genocide researchers. While this statistical model yielded critical insights into the causal factors of genocide, Harff herself recognized the delicate nature of forming policy initiatives based on these findings. Encouraging and aiding developing democracies appears to be an obvious step based on her results; however, the case is not simple. She states that "[a]ttempts to force democratization are problematic because such attempts in poor, heterogeneous countries often fail." At the same time, she found that the "economic connectedness" of countries on a global scale diminishes the likelihood that "...elites will target minorities and political opponents for destruction" [82].

Work continues in this area, with risk assessment models advancing given greater availability of data and improved computing capabilities. As an example, the Political Instability Task Force model employs four independent variables: regime type, infant mortality, conflictridden neighborhood, and state-led discrimination. This very simple framework can predict the risk that a stable country will experience instability in the form of violence or state failure with 80% accuracy [89]. Next, Rost contributes a structural model that shows a high degree of correlation between violent and/or economic repression of marginalized groups and "political threats" like riots and assassinations [90]. The Australian Atrocity Forecasting Project (AFP), headed by Benjamin Goldsmith, developed a forecasting model around structural features including regime type, history of violence and genocide, and general instability that predicts the risk of genocide in one year with 90% and 79% accuracy in positive and negative cases respectively [91]. In their paper "Genocide Forecasting: Past Accuracy and New Forecasts to 2020," Goldsmith & Butcher report continued developments in the work of the AFP including algorithmic refinements to improve predictive accuracy [92]. Additionally, Goldsmith et al. have recently compiled the "Targeted Mass Killings Data Set," which will allow further advancements in this area by contributing a new set of variables and measures useful for understanding and predicting genocide and mass violence [93].

Early Warning Models. As evidenced by the risk assessment models discussed above that yield forecasts of the risk of mass violence and genocide, there is not a fine line that differentiates these from early warning models. In general, early warning models provide shorter term predictions of risk than risk assessment models, and "...include dynamic factors and are meant to be sensitive to changes in political context that trigger violence" [39]. These models are often developed and used by governmental and non-governmental organizations in order to provide timely warnings to policy makers that short-term changes have made violence more probable. Work in this area includes that of the UN Office on Genocide Prevention and the Responsibility to Protect [94], Genocide Watch [95], the International Crisis Group [96], the Sentinel Project [97], and the Global Centre for the Responsibility to Protect [98]. As an example of the contributions of this work, the Global Centre for the Responsibility to Protect shows that populations at highest risk of persecution in the near term include Uighurs in China, and many civilians living in Democratic Republic of the

Congo. This organization advocates for these and other groups at risk at the United Nations, using the results of their predictive algorithms to provide evidence of extreme, near-term risk [98].

Connections, Strengths, and Limitations. The above work highlights the power of computational analysis of data in order to study and predict atrocities like genocide. Risk assessment models highlight structural features of importance between societies, while early warning models can provide much needed alerts to policy makers when shorter term conditions change such that atrocities are potentially imminent. However, as Verdeja is careful to point out, these models are predictive and not causal. Predictive models such as these use data to quantify probabilities of future outcomes, while a causal model attempts to "explain" the phenomenon. Correlation of data to outcomes can provide highly useful predictions; however, it does not necessarily explain causal mechanisms [39].

How can this work inform an ABM? The results of models reviewed in this section can be highly useful to the agent-based modeler, particularly when the goal of an ABM is to be able to reproduce specific scenarios and then perform counterfactual analysis to determine the likelihood of different outcomes. For example, Barbara Harff's work shows that "exclusionary ideologies" are highly significant in predicting genocide, and an ABM of Rwanda might implement that variable in order to be able to use real-world data to inform the model environment [82]. Results like these help the agent-based modeler sift through endless possibilities, zeroing in on those variables that the data show are most informative. It is critical to note that the above models do not directly simulate interaction. In Harff's case, she clearly notes that her analyses illuminate the difficulty of deriving effective policy initiatives from findings through the use of statistical models [82]. The complexity of human systems makes the potential of unintended and unforeseen consequences a reality that cannot be ignored, and agent-based models provide one avenue for rectifying this problem.

3.4.2 Agent-Based Models of Conflict

Overview. The work described above in Section 3.4.1 provides invaluable information useful for identifying structural features that put societies at risk, and also for predicting when atrocities are likely to arise or worsen. These models and results are particularly useful to policy makers working toward peace and stability, both locally and globally. However, as Verdeja notes, they are designed with a predictive purpose in mind, rather than for "explaining causal relations" [39]. Theory-driven, agent-based models can be designed to explore causality, yielding and additional lens through which to view problems. Through the process of validation, discussed above in Section 3.2.2, when ABMs can reproduce the targeted output along the spectrum of validation, their underlying theories and other micro-specifications become "candidate explanations" for the outcome [11]. For the policy maker, a critical strength of an ABM is its ability to explore policy scenarios that have a micro-level, individual impact. Determining the likelihood of different outcomes in a complex adaptive system, such as a society at risk of genocide, is essential to their work and ABMs are one tool in a set of many needed to achieve that goal.

Seminal Models. Epstein's model of civil violence is widely cited and considered a seminal ABM in this area [99]. In this simple model, there are two types of agents: civilians and cops. Civilians have individual attributes measuring their level of "Hardship," perception of governmental "Legitimacy," "Grievance" which is based on the former two attributes, and "Risk Aversion." Epstein's model yielded emergent phenomena like deceptive behaviors and overall breakdowns of the society. These results and others support Epstein's hypothesis that a "...highly idealized model is sufficient to generate recognizable macroscopic revolutionary dynamics of fundamental interest" [99].

Next, Axelrod's model of "The Evolution of Ethnocentrism" is foundational to the area [100]. This model implements a game-theoretic exploration of ethnocentrism. Agents are from one of four identity groups, and decide to help their neighbors based on traits that determine their strategy for cooperating (or not) with neighbors from other groups. Axelrod found that in this simple model, ethnocentric strategies were more robust over a "wide range of parameters and variations in the model." This was an important finding because this preference was not explicitly imposed; rather, it emerged from individual-level, local agent interaction and decision-making [100].

Finally, Bhavnani's model of the effect of ethnic norms on violence in the Rwandan genocide is one of the earliest that explores a specific genocide using an agent-based model [101]. In this simple model, agents have attributes of "animosity" toward other ethnic groups, "tolerance" for those of their same group who have different levels of animosity, and "influence" when interacting with other agents. Agents can be punished for noncompliance with the actions of their ethnic group. Bhavnani found that punishment was required in order for violence promoting norms to develop in the artificial society. Without experiencing this negative consequence for noncompliance, genocidal violence was restrained regardless of the initial makeup of the agent population and conditions of the environment. While the model is not empirically validated on the Rwandan scenario, its reproduction of known patters from the genocide mean that its foundational connection between norm formation and punishment is a possible explanation for mass participation in the genocide.

Recent Models. In this relatively new field, the seminal models above provide frameworks, methodologies, and meaningful results that have driven the next set of questions and advancements. Relevant ABMs in this area not only explore different dimensions of the problem, they have different goals. As shown in Figure 3.1, these goals drive model development and validation. Typically a model of identity and/or civil conflict falls into one of two validation categories: empirical validation, or validation on stylized facts. Models with empirical validation can reproduce the outcomes of a specific scenario, while those validated on stylized facts reproduce general patterns as qualitatively or quantitatively understood. Both approaches have value, as we explain below. Von Briesen et al. review a number of more recent models in the area, clarifying the research gap between models that are validated empirically and those validated based on stylized facts [51]. Next, we summarize main points of these two categories of models.

Empirically validated models. Each of the following models has a goal of reproducing scenario-specific dynamics and achieves empirical validation.

- *"Modeling Civil Violence in Afghanistan: Ethnic Geography, Control and Collaboration,"* Bhavnani & Choi, 2012 [102].
  - **Goal:** Determine the effect of geographic location on the spread of ethnic violence in Afghanistan.
  - **Main result:** Powerful, elite actors are more likely to attack out-groups in a heterogeneous environment; lack of total control of an area increases the likelihood of out-group violence; "concentrated" groups of minorities are at higher risk of violence.
- *"Violence and Ethnic Segregation: A Computational Model Applied to Baghdad,"* Weidmann & Salehyan, 2013 [103].
  - **Goal:** Determine if troop surges or the creation of ethnically homogeneous areas were the cause of decreased violence in Baghdad.
  - **Main result:** Increased segregation leads to a reduction in violence, and early policing action can reduce ethnic violence in polarized regions.
- *"Group Segregation and Urban Violence,"* Bhavnani, et al., 2014 [104].
  - **Goal:** Determine how segregation patterns influence "intergroup violence" in Jerusalem, and use the empirically validated model to explore the potential effects of different policies.
  - **Main result:** Social distance and segregation parameters that allowed for empirical validation sufficiently represented ethnic tensions in Jerusalem. Counterfactual analysis showed that a return to the 1967 borders would have the greatest beneficial impact on violence.

Models validated on stylized facts. Models in this category reproduce a wide range of behavioral dynamics and achieve validation based on stylized facts as understood by subject matter experts.

- *"Ethnic Polarization, Ethnic Salience, and Civil War,"* Bhavnani & Miodownik, 2009 [105].
  - **Goal:** Determine how ethnic salience can "moderate" ethnic polarization and civil war.
  - **Main result:** Salience is a "key moderating variable" with respect to the evolution of polarization and war.
- *"Rumor Dynamics in Ethnic Violence,"* Bhavnani et al., 2009 [87].
  - **Goal:** Determine the effect of rumor dynamics on the spread of ideological extremism.
  - **Main result:** Rumors "survive" in the environment more readily when group leaders are extremists with frequent, low-level social contact.
- *"Micro-cleavages and violence in civil wars: A computational assessment,"* Weidmann, 2016 [106].
  - **Goal:** Determine the effect of micro-level alliances that are supported by macro-level actors on outcomes in civil war.
  - **Main result:** These alliances increase the "severity" of civil war given the macro-level actor supports local alliances and the violence occurs in "rural areas."
- *"How ethnic structure affects civil conflict: A model of endogenous grievance,"* Kustov, 2017 [107]
  - **Goal:** Determine the effect of dimensionality of ethnic identity on civil war.
  - **Main result:** "Crosscuttingness" of ethnicity, in which salience is reduced given a bi-dimensional attribute, reduces overall grievance; however, when violence does emerge, it has increased intensity.
- *"A revolutionary crowd model: Implemented to contrast oscillating to consistent media influence on crowd behavior,"* Ibrahim & Hassan, 2017 [108].
  - **Goal:** Extend prior work in revolutionary crowd models to determine the effects

of different levels and patterns of media intensity on crowd dynamics.

  - **Main result:** The influence of "acquantances" is not sufficient to maintain protest momentum. In addition to consistent use of social media, repeated signaling from recognized leadership is essential to sustain a movement.

- *"A Generative Model of the Mutual Escalation of Anxiety Between Religious Groups,"* Shults et al., 2018 [109].
  - **Goal:** Determine the effect of environmental "hazards" on anxiety experienced between religious groups.
  - **Main result:** Anxiety between groups increases when their overall sizes is similar and agents exceed their anxiety thresholds in the presence of hazards. These conditions allow for more frequent inter-group contact, as well as increased opportunity for agents to perceive members of other groups as "threats" given high levels of anxiety.

3.4.3 Gaps and Implications

One of the strengths of agent-based models is their ability to simulate micro-level heterogeneity and its impact on emergent outcomes over time. In the area of conflict studies, this is often used to explore micro-level parameters that influence societal peace and stability, as detailed in the description of models validated on stylized facts. Empirically validated models like those described above have an additional value in that they can be used to perform counterfactual analysis. Bhavnani et al.'s model of Palestine is a prime example of this capability, as it yielded quantitative results that showed the probability of success or failure of different policy initiatives [104]. However, re-purposing that model to explore a different scenario is a difficult process given its highly specific nature.

Returning to the spectrum of model output validation shown in Figure 3.1, generalized models that are validated on stylized facts can yield results useful for exploring scenario dynamics, such as Bhavnani's model of the Rwandan genocide [101]. Here a very simple model can reproduce genocidal violence in an abstract manner, unveiling the importance of norms, punishment, and other factors, on this emergent outcome. However, the model is not designed to reproduce actual events except in a stylized fashion, which Bhavnani suggests is a problem to address in future work [101]. Alternatively, Bhavnani's later model of Afghanistan is designed to be much more scenario specific, and is empirically validated [102]. This empirical validation allows the model to be used for counterfactual analysis, exploring "what if" scenarios. Since the model can be trusted to reproduce known events, its results for counterfactual analysis are more robust and useful for informing policy and the like.

Research gap. *The gap between these types of models is that the general are not scenario specific, and the scenario specific cannot be easily generalized.* Chapter 4 describes the Eris model, which is designed to be both general and specific. Its theoretical framework is introduced below in Section 3.5. This framework is the basis from which the model will fill that gap, yielding an ABM of identity-based conflict that lies in the "Other Forms of Validation" portion of Figure 3.1's validation spectrum visualization.

3.5 Bystander Motivation During Identity-Based Conflict: A Model Framework

Complexity theory tells us that actions on a micro-scale can lead to unexpected, emergent outcomes. Computational social science provides a generalized framework for implementing computational models of social systems, which are often informed by principles from complexity theory. Returning to the problem of genocide as discussed in Chapter 2, the next step in constructing an agent-based model of this problem is to use the lessons of history and the theories of social scientists in harmony with complexity and CSS in order to develop a model framework.

Macro-level framework. In order to simplify the macro-level model framework, we abstract Straus' factors of restraint as explained in Section 2.2 into a system-level β function. The higher the value of β, the stronger the structural features of a society that prevent persecution of marginalized groups.

Micro-level framework. In order for the theories from Section 2.3 to effectively inform the Eris model, they must yield a framework that connects micro-level states with agent decision-making. Table 2.2 established connections from motivational orientation, to the types of goals and affect associated with that orientation, to a decision-making mechanism that can be implemented in the model. An agent with a high egosystem value will have affective states that make it more likely to decide to be a complicit bystander. In contrast, an agent with high ecosystem value is more likely to be an active bystander, as it is motivated toward compassion and has affective states like love and cooperativeness.

This research seeks to contribute to the research area of theory-driven CSS through the development of a novel, agent-based model of identity-based conflict. The use of data in this model is for the purposes of both informing the environment to explore known cases, and evaluating how closely the model results match historic outcomes. The goal of the work is to provide a generalized model that explores human behavior in the presence of conflict around the issue of identity, seeking to understand how individual motivations and choices impact outcomes.

The above theoretical framework combines the collective knowledge reviewed in this chapter as well as that from Chapter 2. Next, Chapter 4 details the full implementation of Eris, an agent-based model of identity-based conflict. The chapter defines and describes the model environment, agent types, and related attributes. It then details the logic behind interaction in the model, as well as its overall assumptions and limitations.

CHAPTER 4: ERIS – A MODEL OF IDENTITY-BASED CONFLICT

Research Question 1 posed in Section 1.2.1, asks: **What is the optimal, small and efficient, computational model of identity-based conflict, potentially leading to full-scale genocide, that will be useful in providing a means for unveiling and exploring the social dynamics of this problem and how these impact emergent outcomes?** In this chapter, we introduce the Eris model as an answer to that question. The model is named after the Greek goddess of "chaos, strife, and discord" [1]. She is known as the "personification of strife" [110], and thus provides an apt name for this simulation of conflict. However, the model defies the goddess' nature, as its purpose is to understand this type of conflict such that it can be better prevented in the future. This aligns with Eris' opposite, Harmonia, the Greek goddess of "concord and harmony" [111].

The Eris model is constructed for the purpose of closing the research gap identified in Section 3.4.3. This chapter outlines the components and underlying logic of the model. As an introduction, Section 4.1 provides a very high-level overview of the model. Next are sections detailing the model's environment (4.2), agents (4.3), and interaction rules (4.4). Section 4.5 covers assumptions and limitations of the model, and Section 4.6 details the model verification process. Finally, Section 4.7 provides an introduction to and overview of the research experiments conducted with the Eris model.

4.1 Overview

We have explored the problem of genocide through the lens of social science in Chapter 2, developed a theoretical framework in Chapter 3, reviewed relevant computational models in Section 3.4 of that chapter, and are now positioned to construct an agent-based model of identity-based conflict and genocide. Figure 4.1 below is a visualization of the Eris model's major components and is designed to be a point of reference to the reader as he or she learns about the details of the model in later sections. In the figure, there are three types of agents: in-group civilians, perpetrators, and out-group civilians. They each have a role in the environment, with the in-group choosing a bystander role based on their motivations, perpetrators trying to harm the out-group, and the out-group's risk of harm being dependent upon the choices of the in-group. The environment contains a β function that is a global measure of restraint against out-group persecution. Higher levels of β indicate that the society's in-group as a whole will resist out-group persecution, such as in a pluralistic democracy. Lower levels of β indicate that the in-group is less likely to intervene in support of the out-group, such as in a polarized and autocratic society.

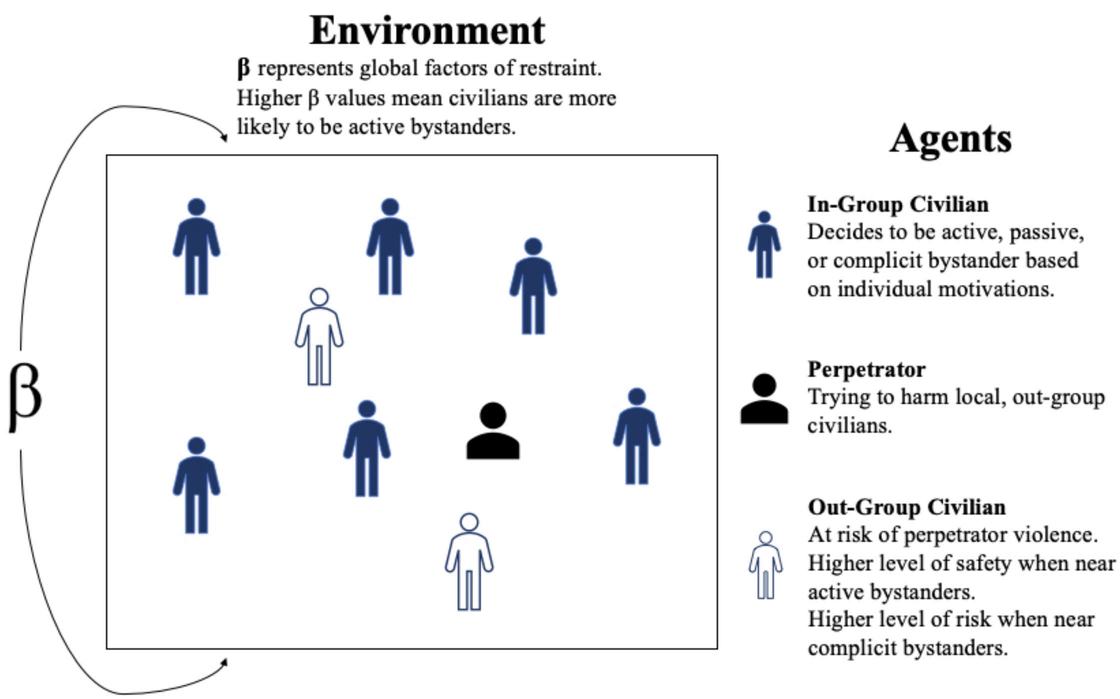


Figure 4.1: High-level visualization of Eris model's major components.

Two major theories from the social sciences inform the attributes and interaction rules of the model components. The theoretical framework detailed in Section 3.5 relies on Crocker & Canevello's "Ecosystem and Egosystem Theory of Motivational Orientations" [10], from Section 2.3.2, to inform the lowest level of the model: agent attributes and interactions. All civilians live in a world where the out-group is at risk if there are perpetrators present, and the base motivations of individual members of the in-group to support and protect the out-group ultimately determine the outcome. High levels of egosystem, relative to ecosystem, will lead an agent to have a higher probability of being a complicit bystander, and the opposite case increases the probability of being an active bystander. The overarching environmental factor, β, is informed by the work of Scott Straus regarding factors of restraint [7], and is defined in Section 2.2. This global factor has a micro-level impact, effectively "pushing" and agent's probability of bystander choice in one direction or the other according to the type of society in which it resides. Together, these components form the the basic framework of the Eris model.

To remind the reader, the overarching goal of this work (see Figure 2.1), is to develop and verify an ABM whose results can influence analysis at any level. The model's foundational theories provide "candidate explanations" [11] for the emergent outcomes, which can then be helpful in refining or establishing new directions in social science research. The Eris model outlined below is based on the work of von Briesen et al. in "Modeling Genocide: An Agent-Based Model of Bystander Motivation and Societal Restraints" [51]. The purpose of changes to that model are to streamline the code, making it more efficient and improving run-time, and also to introduce additional user options to model scenarios based on real-word events. Next, we proceed from high- to low-level in introducing the Eris model, explaining all components and logic in detail.

4.2 Environment

Returning to the discussion of ABMs in Section 3.2.1, the environment in these types of models is simply defined as the world in which its agents reside. There are many possible "worlds," and they are often based on real scenarios. In the case of the Eris model, the environment is a polarized society in which there is a potential for conflict around the issue of identity. Its purpose is to explore the dynamics of this type of conflict, seeking to understand which conditions engender peace, and which make out-group persecution more likely. Next, we introduce the Eris model's environment, explaining exactly what it represents, what it does *not* represent, and its formal attributes.

4.2.1 Understanding the Environment

There are two identity groups in the Eris model: an in-group and an out-group. Agents from both groups move freely and randomly through the world, and are joined by randomly moving perpetrators seeking to harm the out-group. At the system level, there are factors of restraint present and known to all in-group agents. These factors influence an in-group agent's motivation to help protect out-group members. Violence can occur in the environment, and is simulated through the removal of out-group members from the artificial world.

In working toward filling the research gap identified in Section 3.4.3, generalized models are not able to be scenario specific, and the scenario specific cannot be easily generalized, the environment of the Eris model must remain highly abstract. Introducing specific neighborhoods, restrictions on movement, and so on, would make the environmental structure too distinct to easily explore dynamics across scenarios. At the same time, filling the above gap requires an environmental element that can be aligned with a real-world scenario or event. To achieve this, the Eris model revisits Scott Straus' work from Section 2.2, implementing factors of restraint at the macro-level in an customizable manner. The next section provides macro-level implementation details of the model environment.

4.2.2 Attributes of the Environment

Table 4.1 provides an overview of three system-level elements present in the Eris model environment. A more detailed explanation of the of the implementation of factors of restraint and escalation follows, clarifying how each is represented.

Macro-level factors of restraint–β. In the Eris model, β is a global function that represents structural features of a society that restrain the persecution of marginalized groups. Returning to the causal loop diagram in Figure 2.2, it is clear that different variables can restrain in-group compliance in the presence of an extreme elite ideology. The variable of

Table 4.1: Eris model environment attributes and global settings.

Global Attribute/Setting	Explanation
β	Function that represents societal **factors of restraint** against out-group persecution.
Crisis Trigger	The user can define the time-step on which the number of perpetrators increases by a specified amount. This simulates the increase in subnational or other actors seeking to harm the out-group on behalf of the societal elite, and is an abstraction of **factors of escalation**.
Temporal Scale	One time-step in the model represents a 24-hour day.

interest here is "Preexisting Democratic or Inclusive Ideals." The model simulates the effect of these ideals over long periods of time to determine how well they can restrain violence against an out-group should escalating factors like war or other threats emerge. Of particular interest in this research is how these ideals are communicated from the societal elite to the civilian population, as quantifying this can provide an avenue for model validation. The use of data to inform the model's β function, and thus achieve empirical model validation, is explained in Chapter 5.

In the model, $\beta \in [0.0, 1.0]$, and can be constant or changing in time according to a user-defined function. $\beta = 0.0$ represents a highly authoritarian society with low levels of restraint and inclusiveness, and $\beta = 1.0$ represents a highly pluralistic society which tolerates no violence against, or persecution of, out-groups. The current version of Eris only implements β as a constant in time, or β as a positive or negative step function.

Figure 4.2 visualizes examples of a constant β at the extremes, as well as possible β step functions where its level rises or falls suddenly on a particular day. A constant β value represents a society in which the elite deliver a consistent message over time with respect to their ideals; the society does not become significantly more or less tolerant of out-groups over time. β represented by a step function simulates a sudden shift in the message from the elite to the society. The change can be large or small, positive or negative. In these cases, this represents a society that has experienced a change influencing the ideals the elite are communicating to its citizens. A positive step could represent the election of a new, highly inclusive president. A negative step could represent internal unrest of some kind that causes the elite power structure feel threatened.

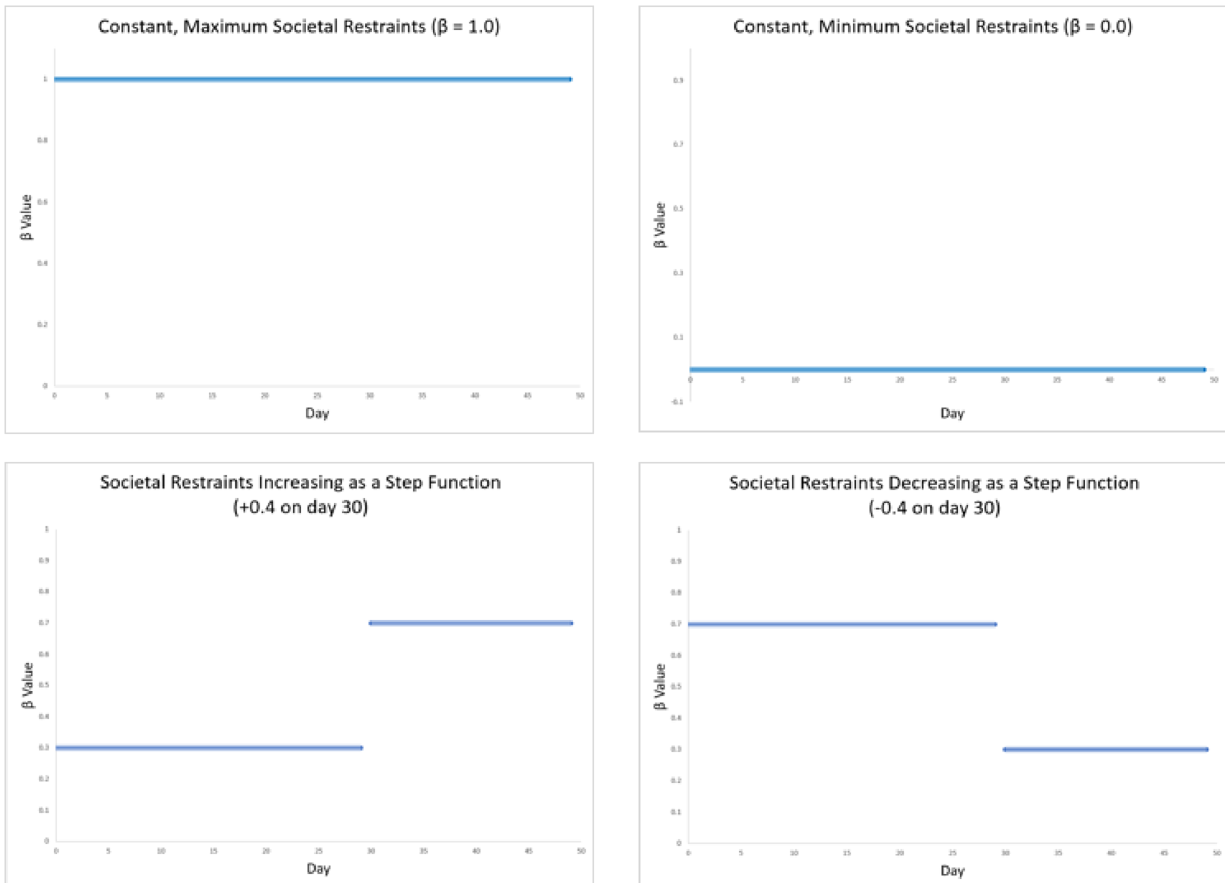


Figure 4.2: Examples of β function values in Eris model.

It is important to note that the above are only *two* highly simplified representations of factors of restraint that are currently implemented in the Eris model. These are the simplest abstractions of β, and serve as the starting point for the model. Future work will include a variety of options for this function, including linear, exponential, or a function informed by custom data input.

Macro-level factors of escalation. The current version of the Eris model does not implement a global function that abstracts factors of escalation in the way β does with factors of restraint. To simplify the model, we assume that low levels of β are sufficient to represent societies in which out-group persecution is more probable. As shown in Table 4.1, to simulate real-world events that may trigger a genocide, the model implements a "Crisis Trigger" setting. To allow a triggering event to occur in the simulation, the user can specify the day on which this event occurs and how many perpetrators should be added to the environment at that time. The increase in perpetrator count can be understood as an abstract representation of the elite drawing on subnational support, as shown in Figure 2.2.

The above completes the description of the Eris model implementation at the system level. The "world" that it represents is highly abstract, and is focused on restraining and escalating factors present in the environment. Next, we define the agents that reside in this environment. Their role in the simulation is critical, as it is their choice of action that ultimately determines if there will be peace or violence in their world.

4.3 Agents

As explained in Section 3.2.1, an agent in an ABM is a discrete and autonomous construct. Agents in the Eris model represent three categories: in-group civilians, out-group civilians, and perpetrators. Perpetrators seek to commit acts of violence against out-group civilians, and require the compliance of the in-group to achieve their goals. In general, perpetrators represent sub-national actors as shown in Figure 2.2, acting in behalf of an elite regime.

4.3.1 In-Group Civilian Agents

The "in-group" defined. In the Eris model, the in-group represents a dominant population with the same identity as the elite leadership in the society. This group is not required to have a numerical majority, as "subaltern" groups have also committed genocide on rare occasions [19]. The model allows the out-group population percentage to be $0\% <=$ (out-group population) $<= 100\%$ in order to accommodate a wide variety of population distributions. More specifically, in-group agents in the model are civilian members of the society. They are not members of the elite, or of their sub-national supporters. As a result, they themselves do not have the power to initiate or stop violence against the out-group; rather, their power lies within their actions as bystanders.

In-group agent characteristics. In-group agents in the model are civilian bystanders witnessing persecution of the out-group by perpetrators working to further an extreme elite agenda. Returning to Table 2.1, Staub defines three types of bystanders: active, passive, and complicit [54]. Active bystanders work to resist, protest, or otherwise prevent persecution of the out-group. Passive bystanders choose not to act in behalf of that group, and complicit bystanders are known by perpetrators to support their agenda.

Importance of in-group civilian bystanders. Staub finds that the actions of bystanders influence the outcome as a society moves toward genocide [52, p. 5]. As shown in in Figure 2.2, a compliant in-group civilian population is necessary in order for an extreme elite agenda to be reinforced, thus allowing genocide to occur. Active bystanders work directly toward restraining this process, and can serve to protect the out-group.

In-group agent attributes. Table 4.2 outlines the in-group agent attributes implemented in the Eris model. This set of attributes is taken from the work of von Briesen, et al. [51], and continues to serve as an agent-level framework for the Eris model.

Note that the last entry in the table is not a formal attribute; rather, it is a derivative of two other attributes. OrientationDifference is used to determine the likelihood that an agent will choose a particular bystander type, and is computed according to the following formula:

$$OrientationDifference = (Egosystem - Ecosystem) \tag{4.1}$$

A positive OrientationDifference means that agent has a higher probability of becoming a complicit bystander, and a negative value means the agent has a higher probability of becoming an active bystander. The reasoning behind this logic is outlined in Table 2.2.

Finally, for attributes initialized from random normal distributions [0.0, 1.0] there is a very small chance of an extreme outlier falling just outside of that range. This model is implemented in NetLogo, which does not have a built in option for a bounded normal distribution [71]. These outliers are extremely rare due to the relatively small standard

Table 4.2: Civilian agent attributes as implemented in the work of von Briesen, et al and carried forward to the Eris model. Table adapted from [51].

Attribute Name	**Type**	**Range**	**Initialization**
Identity (ID)	static, binary	ID $\in$ {A, B}	user determines number of agents per identity group
Susceptibility (S)	static, float	S $\in$ [0.0, 1.0]	normal distribution ($\mu = 0.5$, $\sigma = 0.16$)
Egosystem (EG)	dynamic, float	EG $\in$ [0.0, 1.0]	normal distribution ($\mu = 0.5$, $\sigma = 0.16$)
Ecosystem (EC)	dynamic, float	EC $\in$ [0.0, 1.0]	normal distribution ($\mu = 0.5$, $\sigma = 0.16$)
Influence (INF)	static, float	INF $\in$ [0.0, 1.0]	normal distribution ($\mu = 0.5$, $\sigma = 0.16$)
BystanderType (bType)	dynamic, ternary	bType $\in$ {-1, 0, 1}	1 = active bystander 0 = passive bystander -1 = complicit bystander
OrientationDifference (OD)	derivative of EG and EC attributes	OD $\in$ [-1.0, 1.0]	$OrientationDifference = Egosystem - Ecosystem$

deviation around the mean, and they do not impact the results. Any variation they introduce is averaged out through the use of 100 replicates to determine final results (see Section 4.6.3), and also does not affect agent decision making due to the bounding of OrientationDifference between [-1.0, 1.0] as clarified in the introductory paragraphs of Section 4.4. Additionally, these outliers can be taken to simulate the appearance of individuals at the very extremes of society.

Micro-level factor of restraint – Active Bystander Contagion. At the micro-level, the Eris model retains an "Active Bystander Contagion" effect as implemented by von Briesen et al. [51]. This is based on Crocker and Canevello's finding that compassionate goals "foster trust" [10]. The model framework detailed in Table 2.2 shows the logic behind higher ecosystem levels leading to a higher probability of an agent becoming an active bystander due to individual affective states. In their model, von Briesen et al. allow an agent with active bystander neighbors to move its OrientationDifference in the direction of the average of these neighbors, thus causing the adapting agent to have a higher probability of becoming an active bystander itself [51].

Micro-level factor of escalation – In-Group Fear. As implemented by von Briesen et al. [51], the Eris model interface includes an optional "In-Group Fear" effect. When this effect is used in a simulation, in-group agents become more afraid if they are near places where there has been past violence. While the system itself does not adapt in any way other than to keep a record of acts violence and their specific locations, this record has a micro-level impact. When in the vicinity of locations on which there has been violence, an agent has an increased likelihood of becoming a complicit bystander due to fear, which serves as a micro-level factor of escalation.

4.3.2 Out-Group Civilian Agents

The "out-group" defined. The out-group in the Eris model are civilian agents that are *not* members of the in-group. This is a simplification of what an individual considers to be his or her identity, as this is often multi-dimensional and can be modeled as such [107]. However, one of the goals of this work is to provide a small and efficient model of genocide, which can be best achieved by allowing for only two identity groups. Like the in-group, out-group agents in the model are civilians. Again, this is a simplification, as persecuted peoples have been known to actively work to defend themselves (see Rwanda discussion in Section 2.1.3).

Characteristics of out-group agents. Out-group agents in the Eris model are at the mercy of local perpetrators unless they are protected by in-group active bystanders. Again, this is a highly simplified representation of the reality of persecuted peoples, as they may resist, collaborate with perpetrators, and so on. Returning to the goal of an optimal, small and efficient model, this is a necessary starting point. Future research as outlined in Section 8.3 provides a road map for model refinements that give the out-group increased dimensionality and agency in their artificial society.

Out-group significance. A genocide cannot occur without an out-group. In an ideal society, there are no factions, no divisions, no recognition that differences in religion, ethnicity, or any other marker of identity gives superiority to any one group over others. However, as history has shown, the existence, marginalization, and persecution of out-groups is a reality that we as a global community must continue to acknowledge.

Out-group agent attributes. In the current version of the Eris model, out-group agents have only one attribute: Identity. The type, range, and initialization for this attribute are identical to what is detailed in Table 4.2. Here, an out-group agent's Identity attribute is simply set as the opposite of the in-group Identity value.

4.3.3 Perpetrator Agents

Perpetrator agents defined. Perpetrators in the model are an abstraction of sub-national actors shown in Figure 2.2. They support the elite agenda, and target for persecution any out-group member in their local neighborhood. Their ability to achieve their goals depends on the ratio of complicit to active bystanders in their local neighborhood; higher ratios of complicit to active bystanders means perpetrators are more likely to successfully persecute the out-group.

Characteristics of perpetrators. Perpetrators in the Eris model are an agent type that is distinct from in-group and out-group civilian agents. While they act to support an extreme elite agenda, and thus represent the interests of the in-group elite, they cannot become civilian agents, and civilian agents cannot become perpetrators. This is a model simplification and limitation that is left as an area of future work outlined in Section 8.3.

Significance of perpetrators. Perpetrators are essential to the problem, as they are the agents of violence against the out-group. Section 2.1.2 describes the nature of perpetrators, their motivations, and their role in the evolution and execution of genocidal violence.

Perpetrator attributes. Perpetrators in the Eris model are the most simple agents. Other than their agent type, they have no other attributes. Their only role in the simulation is to move randomly and threaten the existence of any out-group agent in their local neighborhood.

4.3.4 Additional Global Settings Affecting Agents and Their Interactions

In addition to the above agent framework, the Eris model includes five global settings that affect agents. Two affect all agents, and three affect only the in-group. These are taken from von Briesen et al. [51].

RadiusofSight (all agents). This is a globally set distance that defines what constitutes a "local" neighborhood. The user determines if this will be a small or large radius, and once set, it is the same for all agent types throughout the simulation.

Movement (all agents). All agents in the simulation move randomly. Their local neighborhood is defined by their location at any given time-step as measured by their RadiusofSight.

ProbabilityofMutation (in-group agents). This is the probability that any in-group agent will randomly reset its attributes. In the absence of the potential for random mutation, in-group agent Egosystem and Ecosystem attributes will converge to the mean, which does not correctly simulate human societies. To address this issue, Bedner, et al. introduce a "behavioral tremble" into their agent-based model, which allows realistic simulation of what they find as "persistent heterogeneity" in cultures [112]. The Eris model follows the work of these authors by allowing for this probability of mutation.

ProbabilityofDeath (in-group agents). This is the probability that any in-group agent will die in that time-step. When an agent dies, it automatically produces one offspring for replacement. The offspring only inherits the parent identity, with all other attributes being set randomly according to the distributions detailed in Table 4.2. This is a highly simplified representation of natural death. It can be customized to represent real-world death rates in particular societies or left as a generalized value.

SusceptibilityFraction (in-group agents). This fractional value is used to slow the rate of change as in-group agents update their attributes during interactions with other in-group agents as detailed in Section 4.4, Algorithm 1. This value can be made higher or lower depending on the user preference for rates of adaptation, with the optimal setting allowing the distributions of Egosystem and Ecosystem across the in-group population to remain stable.

This concludes the overview of the Eris model's environment and agent framework. Next, we define the rules for interaction in the model, connecting all of the above in order to simulate the dynamics of identity-based conflict.

4.4 Interaction Rules in the Base Model

Before proceeding to the details of Eris model interaction, it is important to clarify that this is an "Adaptive ABM" as defined in Table 3.1. This is different from other types of ABMs listed in the table, as it contains agents that change in time as they interact. Agents move randomly through the environment. In-group agents interact and adapt according to their local in-group neighbors as well as the system-level conditions implemented in the β function. Only in-group agents make decisions, and their decisions determine their bystander types. Out-group agents are passive, and perpetrators are seeking to attack any local out-group agents. Their failure or success rests on the decisions of in-group bystanders.

The Eris model maintains the general logic implemented by von Briesen et al. [51], and includes several modifications:

1. In-group agent susceptibility mitigates adjustments to OrientationDifference due to the presence of:
   (a) β function (Algorithm 2)
   (b) Active Bystander Contagion effect (Algorithm 3)
   (c) In-Group Fear effect (Algorithm 4)
2. Macro-level factors of escalation, β, can be constant in time or a step function (see Section 4.2.2).
3. The user can specify a CrisisTrigger as a macro-level factor of escalation. On the chosen day, perpetrators will increase in number according to the user-defined setting (see Section 4.2.2).

For logic outlined below that impacts an agent's OrientationDifference, note that the underlying code is adjusted to ensure that $OrientationDifference \in [-1.0, 1.0]$. It is mathematically possible for the relevant factors to push OrientationDifference outside of this range if an agent is heavily weighted in the direction of Egosystem or Ecosystem. This logic is different than what appears in Algorithm 1 during local adaptation, where one in-group agent moves its personality attributes in the direction of another, more influential agent. Here, factors such as β in Algorithm 2, Active Bystander Contagion in Algorithm 3, and In-Group Fear in Algorithm 4 have the same effect on all agents regardless of their initial OrientationDifference, mitigated only by an individual agent's Susceptibility to change. These factors push an agent's OrientationDifference in a positive or negative direction regardless of its starting value, which can potentially cause it to fall outside of the $[-1.0, 1.0]$ range.

4.4.1 In-Group Local Adaptation

As an in-group civilian agent moves randomly through the model environment, it randomly selects a neighboring in-group agent at each time step. If that neighboring agent is more influential, the adapting agent will update its Egosystem and Ecosystem attributes in the direction of its neighbor. This interaction rule simulates the ability of people to influence the motivations of others with whom they interact [10, 40].

Algorithm 1 In-Group Adaptation According to Agent Influence (adapted from [51])

let $i = \textit{adapting agent}$
let $j = \textit{random in-group neighbor agent}$
if $Influence_j > Influence_i$ **then**
  $Egosystem_{i(new)} \leftarrow [Egosystem_{i(old)} +$
    $\hookrightarrow (Egosystem_j - Egosystem_{i(old)})] * \frac{Susceptibility_i}{SusceptibilityFraction}$
  $Ecosystem_{i(new)} \leftarrow [Ecosystem_{i(old)} +$
    $\hookrightarrow (Ecosystem_j - Ecosystem_{i(old)})] * \frac{Susceptibility_i}{SusceptibilityFraction}$
end if

4.4.2 In-Group BystanderType Selection

Next, an in-group agent will calculate its OrientationDifference from its Egosystem and Ecosystem attributes. This derivative value is used to determine the probability of the agents choosing to be a passive, complicit, or active bystander. In addition to the micro-level attributes contributing to the BystanderType decision, agents can also account for **macro-level factors of restraint** if the simulation includes a β function. Here, β pushes an agent's OrientationDifference in the negative direction according to the its susceptibility to change, increasing the likelihood that the agent will choose to be an active bystander. Note that β is determined by the model user, and can be a constant, a step function, or zero if he or she desires to have no global factors of restraint present.

Algorithm 2 In-Group Agent Determination of BystanderType [51]

let $OrientationDifference = Egosystem - Ecosystem$
if $\beta = TRUE$ {There is a β function in the system} **then**
  $OrientationDifference \leftarrow OrientationDifference - (\beta * Susceptibility)$
end if
if $OrientationDifference = 0$ **then**
  Agent is a **passive** bystander
end if
if $OrientationDifference > 0$ **then**
  Agent probability to become a **complicit** bystander is determined according to the magnitude of the OrientationDifference. Higher values yield a higher probability.
end if
if $OrientationDifference < 0$ **then**
  Agent probability to become an **active** bystander is determined according to the magnitude of the OrientationDifference. Higher values yield a higher probability.
end if

4.4.3 Active Bystander Contagion Effect

As explained in Section 4.3.1, in-group agents can experience a **micro-level factor of restraint** if they have active bystanders in their local neighborhood. First, the agent itself must have an OrientationDifference that gives it a probability of becoming an active bystander. In this case, that agent's probability increases according to the average of the

OrientationDifference of its active bystander neighbors. Note that this effect can be turned on or off by the user.

Algorithm 3 Active Bystander Contagion (adapted from [51])

---

let $averageOD$ = *average OrientationDifference of active bystander neighbors*
if $OrientationDifference < 0$ {an agent has a probability to become an active bystander} **then**
  $OrientationDifference_{new} \leftarrow OrientationDifference_{old} -$
    $\hookrightarrow$ $(averageOD * Susceptibility)$
end if

4.4.4 In-Group Fear Effect

As explained in Section 4.3.1, in-group agents can experience a **micro-level factor of escalation** if they are in an area where violence has occurred in the past. Fear due to local violence causes the agent's OrientationDifference to move in the positive direction, making it to become more likely to become a complicit bystander [52, p. 18]. Again, this effect can be turned on or off by the user.

Algorithm 4 In-Group Fear [51]

---

let $totalDeaths$ = *sum of deathCount from all locations within agent's vision*
if $totalDeaths > 0$ **then**
  $OrientationDifference_{new} \leftarrow OrientationDifference_{old} + \frac{totalDeaths}{100}$
end if

4.4.5 Out-Group Violence

Recall that as in-group agents are moving randomly and adapting according to their local and global conditions, perpetrators and out-group agents are also moving randomly. Algorithm 5 details the logic behind the survival of an out-group agent at a given time step and location. First, in order for an out-group agent to be at risk, there must be at least one perpetrator in its vicinity. Next, there must be a greater number of complicit than active bystanders, with a higher proportion of complicit bystanders yielding a higher likelihood of death.

Algorithm 5 Violence Against Out-Group (adapted from [51])

let $totalBystCt$ = *total number of all in-group civilians in local radius*
let $activeBystCt$ = *number of in-group active bystanders in local radius*
let $complicitBystCt$ = *number of in-group complicit bystanders in local radius*
let $perpetratorCt$ = *number of perpetrators in local radius*
if $(perpetratorCt > 0) AND (complicitBystCt > activeBystCt)$ **then**
$probabilityofDeath \leftarrow (\frac{complicitBystCt}{totalBystCt} - \frac{activeBystCt}{totalBystCt})/10$
{probability is reduced by a factor of 10 in order to allow for longer model runs}
end if

4.5 Limitations and Key Assumptions

Translating human to computational behavior is abstract by its nature, and simplification of that process is necessary in order to avoid producing overly complex models whose results are difficult to interpret. As such, a computational model is always limited in scope, and it is critical to clearly understand the limitations of the model and any key assumptions made in determining its framework. In this section, we clarify the most critical limitations and assumptions of the Eris model.

Salience and dimensionality of identity. The Eris model has two major limitations with respect to identity. First, the model includes only two identity groups. This means that it is not appropriate for simulating scenarios in which there is conflict between a larger set of identity groups, as was the case during the Bosnian genocide in the early to mid-1990s. Next, the model assumes that identity is a static and salient agent attribute. Agents are assigned to an in-group or an out-group at the beginning of the simulation, cannot change groups, and their identity is always salient. This is a drastic simplification of how identity is modeled in some prior work, where it has been given variable salience and dimensions [105, 102, 107]. The purpose of allowing these limitations is to simplify the exploration of the dynamics of identity-based conflict without prematurely complicating the issue of identity. However, it is likely that introducing dynamic salience and evolution of identity will be implemented in future work.

Abstraction of violence. What does it mean to "kill" an agent? When an out-group agent is targeted by a perpetrator, the out-group agent simply disappears from the environment. The assumption here is that this sufficiently simulates imprisonment, exile, death, or other forms of persecution, as each leads to that out-group member being effectively "removed" from society. With respect to the acting perpetrator, this also restricts the definition of its action to simple "removal" of an out-group agent.

Out-group adaptation. Next, the Eris model limits adaptation to in-group agents only. Out-group agents are limited to the role of a "helpless victim" in the event of violence; they are in a subservient position with no power to influence their circumstances. Instead, it is up to in-group bystanders to determine the outcome. Future work will allow out-group agents to adapt, resist, and influence the in-group; however, this research focuses on the role of in-group civilians as bystanders, which can be more clearly explored by implementing the model with the above limitation.

Perpetrators as a static group. In the current version of the model, in-group civilians cannot become perpetrators, and perpetrators cannot become civilians. The most an in-group agent can do to support persecution of the out-group is to become a complicit bystander, providing real or perceived support to perpetrators. While this simplifies the nature of the groups within the simulation, it prevents more realistic scenarios in which civilians join in the violence, crossing the threshold from bystander to perpetrator.

In-group motivations. A critical assumption made in the Eris model is that an agent's Egosystem motivation toward "self-image" goals applies to individual as well as the group with which it identifies. A high Egosystem value makes an agent more likely to be a complicit bystander, acting in harmony with the goals of perpetrators and the extreme elite they represent. The assumption is that in the presence of strong salience of identity, an individual will see the interests of the group as equivalent to the interests of the "self." Section 2.3.1 presents social psychology theories from Staub and Swann et al. that support this logic in the model [9, 48].

Death and reproduction. As detailed in Sections 4.3.4 and 4.6.2, the Eris model implements a ProbabilityofDeath of 1 in 25,000 for any in-group agent on any day, with the dying agent producing one offspring with randomized traits at that time. This is an extreme simplification of real-world death and reproduction. Future work with the model should allow for more accurate death and reproduction based on data from the types of societies being explored. The death rate in the current version can be easily adjusted, while more complex reproduction options will require adjustments and additions to the code.

This completes information about the limitations and assumptions of the Eris model. The next section explains the verification of the model, as well as the approach to validation. These are the final elements required before proceeding to use of the Eris model to study the dynamics of identity-based conflict.

4.6 Model Verification

This section covers the verification and validation of the Eris model. Verification is the process by which the model is ensured to be operating without error and as intended. Here we discuss improvements to the model code of von Briesen et al. [51], necessary adjustments to obtain stability in agent attributes as they interact and adapt, and determination of the number of replicates necessary to obtain stable average results.

4.6.1 Improving Efficiency of the Model Code

Code review and debugging are essential to model verification, and this process led to an effort to improve the efficiency of von Briesen et al.'s model [51]. As discussed in Section 1.2, one of the goals of this research is to provide a model that will be useful to social scientists, policy makers, and others. As such, this model should be optimized for use on a local computer. Cloud-based simulation runs will be required if the user wishes to model real-word population sizes; however, at this stage of the work, the goal is to examine patterns, death-rates, and more general aspects of the problem, which is possible in a local computing environment.

Railsback et al. [113] and Grider [114] describe various approaches to making agent-based models more efficient, which motivated an exploration of areas in which the Eris model could be improved. The current version of the model adopts the vast majority of von Briesen, et al.'s framework [51], and implements these two changes:

1. **Removed out-group violence choice for user.** In order to explore a scenario with no violence, the user can choose to have no perpetrators in the environment.
2. **Removed all out-group attributes except for identity.** In the current model, the out-group is passive and does not adapt, which means the model should not allocate memory for these unused attributes. Future versions of Eris will need to reintroduce these attributes in order to explore out-group dynamic change.

The first change streamlines the logic, making the code run more efficiently. The second change reduces the amount of memory required for agent attributes by removing those that are unused.

4.6.2 Stability of In-Group Ecosystem and Egosystem Attribute Distributions

Returning to the research methodology shown in Figure 2.1, the next step toward properly verifying the Eris model applies to intra-group dynamics. Determining the behavior of the model when only the in-group is interacting with itself is essential, as it should show that parameters and results are stable before adding in layers of behavioral complexity. This ensures the validity of results of inter-group dynamics, which allow for violence to occur against out-group members as the in-group interacts and adapts to changing local and global conditions.

Given the adaptation of in-group agents as they interact (see Algorithm 1), it is necessary to account for the tendency of an ABM to converge. This occurs because of the proportionally large number of agents with attribute values close to the mean. Their high level of representation in the environment creates an increased probability of any other agent adapting in their direction. A model that allows this to occur is not a realistic representation of society, where general distributions of traits should remain stable. The distribution should remain normal, continue to have outliers, and have a stable standard deviation.

As shown in Table 4.2, an in-group agent's Egosystem and Ecosystem attributes are dynamic. They change according to the logic outlined in Algorithm 1, and can converge to the mean if there are no moderating factors that ensure diversity in the population. In order to achieve this level of stability, we have included three global model settings allowing for mutation, death and reproduction, and fractional susceptibility. Each is listed and explained in Section 4.3.4. ProbabilityofMutation accounts for the probability that any agent will randomly change, which ensures heterogeneity in the environment [112]. Death and reproduction, the ProbabilityofDeath, also allows for heterogeneity in a system, as offspring are not exact replicates of their parents in the Eris model. Once accounting for the above, the model still requires an additional variable to slow the rate of change overall. This fractional value, the SusceptibilityFraction, reduces the magnitude of change to Ecosystem and Egosystem during adaptation in Algorithm 1, allowing the model that adaptation to occur over a realistic time-frame.

Our exploration and analysis of the model yielded the following settings for these parameters in order to ensure model stability:

- ProbabilityofMutation = 1 in 10,000
- ProbabilityofDeath = 1 in 25,000 [1]
- SusceptibilityFraction = 10,000

4.6.3 Stability of Replicate Averages

Agent-based models like Eris have great deal of inherent stochasticity. This is the result of agent attributes being drawn from random distributions, agent locations and movement being random, and so on. As a result, the user should not rely on the results of a single simulation run to draw conclusions because this outcome may be difficult to replicate. In order to account for this effect, agent-based modelers run replicates of simulations and take the average of the results to determine reliable outcomes. The number of replicates required

[1] Note that the ProbabilityofDeath should be customized for specific, real-world scenarios in order to more accurately model their actual death rates.

is dependent upon the level of model stochasticity and the intended use of the results. For example, a model that will inform official policy decisions has a higher standard than one used for simple problem exploration in early research phases.

In order to determine the number of replicates required to produce stable results in the Eris model, we performed three sets of experiments. The sets tested the results of 10, 50, and 100 replicates on the the out-group population over time. Each experiment had the following settings:

- Number of agents = 500
- In-group percentage = 70% (of total number of agents)
- $\beta = 0.09$
- No Active Bystander Contagion or In-Group Fear effects

Figure 4.3 below shows the results of these experimental sets. Note that in each chart in the figure, there are three separate plot lines that show the results of three different model runs averaged over that number of replicates. For example, the top left chart displays the average out-group count for 10 replicates during that set of runs, repeated three times.

Included in this figure are the standard deviation between results for each set. This is calculated from the standard deviation in average out-group count over days 0-7500. The results are as follows:

- 10 replicates: SD = 2.38
- 50 replicates: SD = 1.45
- 100 replicates: SD = 0.29

The final result, for 100 replicates, has a standard deviation of only 0.29. This means that the mean out-group count taken over 100 replicates will vary by less than one agent on a given day. As such, we have elected to perform all Eris model experiments over 100 replicates except where noted, taking the average of those as the final result.

Finally, it is important to note that these experiments were conducted *without* agent susceptibility moderating the value for OrientationDifference as shown in Algorithm 2. Chapter

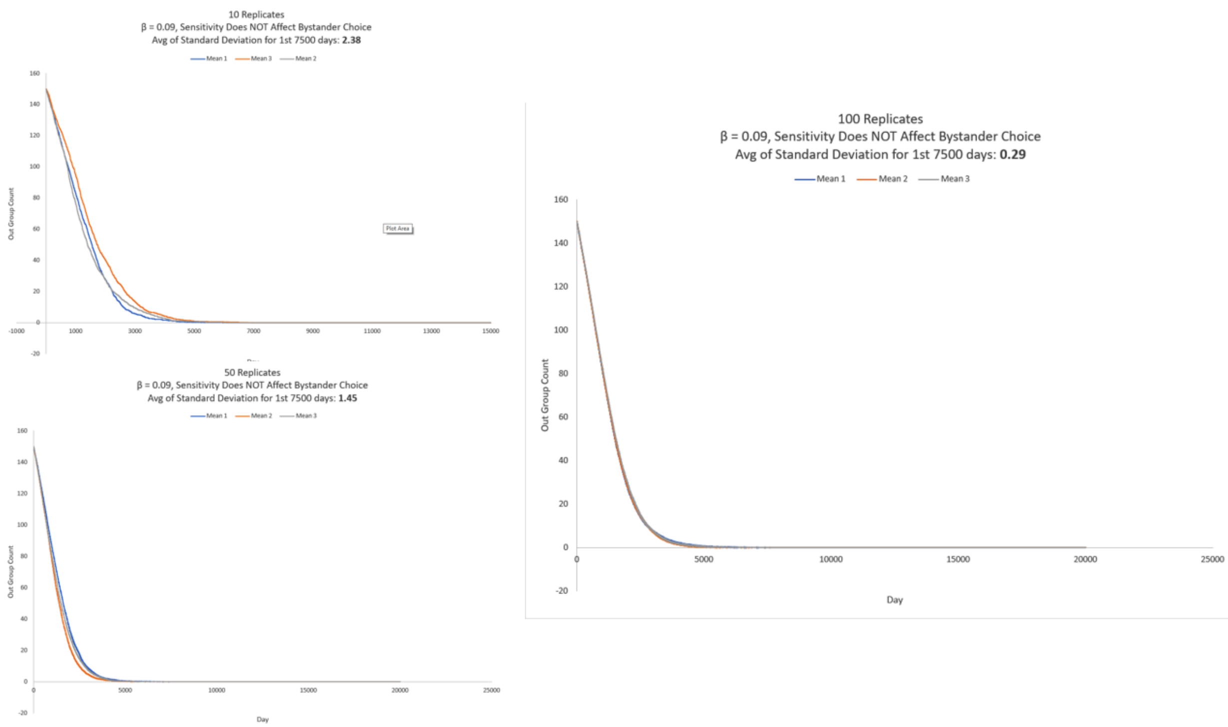


Figure 4.3: Results from simulations to determine number of replicates required to ensure result stability.

6 highlights the difference in results both with and without the moderating effect of Susceptibility on various outcomes, and it is apparent that allowing for Susceptibility to moderate change increases the stochasticity of the results. However, we have elected to use 100 replicates for all results given it is sufficiently stable on average.

We close this chapter with an overview of the experimental process for the Eris model. The goal of this methodology is to successfully validate the model across two different real-world scenarios. This requires appropriate data, a full exploration of the model in a most general sense, and finally, customization for validation experiments.

4.7 Model Experiments and Validation

The next three chapters lead from the above description of the Eris model to its use for empirical validation across real-world scenarios. As discussed in Section 3.2.2, model validation is essential in that it tells the user how well a model corresponds to reality. The level of validation required depends on the goals of the model, and is best understood as a spectrum as shown in Figure 3.1. The Eris model seeks to fill the research gap noted in Section 3.4.3: generalized models of conflict are not easily customized to simulate a real-world scenario, and scenario-specific models lack the flexibility to adapt to different events.

In order to achieve empirical validation, an ABM typically uses data related to the event in question to inform the model environment [102, 103]. For the Eris model, we have chosen to reproduce two contrasting scenarios: the positive case of genocide in Rwanda in 1994, and the negative case in Côte d'Ivoire in the 2000s. The reasoning behind selecting these two cases is covered in Sections 2.1.3, 2.1.4, and 2.2. Chapter 5 covers the data sources used for model validation, analysis of these data, and presents a framework for how they can be used in the validation process.

Chapter 6 covers Experimental Set 1, which explores generalized inter-group dynamics in the Eris model, with model validation based on the input of subject matter experts. These results provide a broad overview of the functionality of the model in exploring problem dynamics. This is an essential precursor to Chapter 7, Experimental Set 2. This set of experiments details the *empirical* validation of the Eris model on the Rwandan and Ivorian cases, and the results presented in Chapter 6 provide context for understanding the model's behavior during the empirical validation process.

This concludes the description of the Eris model framework, its limitations and key assumptions, model verification, and introduction to model experiments and validation. Eris is a model that is grounded in fundamental human motivations for self- and species-preservation. These drive human behavior at a base level, yielding a model that is small and efficient. The following three chapters put that framework in motion, working further toward answering the research questions stated in Section 1.2 regarding data required for validation, and evaluation of its usefulness to a wide range of users.

CHAPTER 5: DATA

The Eris model presented in Chapter 4 is a generalized framework from which to study the dynamics of identity-based conflict. However, one of the goals of this work is to fill the research gap identified in Section 3.4.3: general models are not scenario specific, and scenario specific models cannot be easily generalized. To fill this gap, the Eris model must be able to reproduce the dynamics of different events. As such, the model targets "Other Forms of Validation" on the validation spectrum shown in Figure 3.1, as its goal is to find a balance between the general and the specific.

This chapter addresses Research Question 2 (Section 1.2.2): **What are the most useful data to inform and empirically validate a generalized, agent-based model of identity-based conflict?** We begin with an overview of why and how sentiment analysis will be used in the process of empirical validation (Section 5.1), followed by an overview and description of the selected data (Section 5.2), and finally by a presentation of the results of the analysis, along with a framework for model validation (Section 5.3).

5.1 Data and Eris Model Validation

Working toward empirical validation of the Eris model, we return our focus to the research of Scott Straus detailed in Section 2.2. Straus has extensively examined both positive and negative cases of genocide in Sub-Saharan Africa. His theory of genocide highlights the importance of factors of escalation and restraint in determining positive outcomes, where genocide occurred as expected, and negative outcomes, where genocide was expected and did not occur [7, 8].

As discussed in Section 4.2.2, the Eris model includes a system-level abstraction of factors of restraint that work in conjunction with agents' individual motivations to determine their behavioral choices within that environment (see Table 4.1). In identifying data appropriate for model validation, it is essential to work in harmony with the model framework. There are a wide variety of positive cases of genocide for which one can find empirical data to help with model validation; however, according to Straus, the same level of attention has not been given to negative cases, making it more difficult to find useful quantitative measures of these scenarios [7, 8]. The key here is to connect an aggregate outcome with individual level behavior, and to find data that makes that connection more robust. The right connection will make it possible to flexibly simulate both positive and negative genocidal outcomes.

Given the adoption of Straus' theory of factors of restraint in the model framework, we will focus validation attempts on two particular cases he has explored in depth: Rwanda in 1994 as a positive case, and Côte d'Ivoire in the 2000s as a negative case. From here, we rely on Straus' database of presidential speeches from the Sub-Saharan cases he examined [12]. These speeches can be used to analyze the sentiment and emotional content of elite discourse prior to the outbreak of violence. Sentiment analysis on in-group, elite communications allows for a direct connection between the micro-level theory of Motivational Orientation and macro-level sentiment, as expressed through text, of the in-group and its representatives. Before taking a closer look at the selected cases and their speech data, we provide a brief review of sentiment analysis research.

5.1.1 Sentiment Analysis Literature

According to Yadollahi et al., sentiment analysis falls into two main categories: opinion mining, and emotion mining [115]. This research focuses on emotion mining, particularly the detection and measuring of emotion in text in order to quantify the motivations expressed by elite actors as they relate to their motivational orientation (see Section 2.3.2). There are a variety of techniques and algorithms for determining emotion in text, and we will focus on the use of a lexicon based approach in this research.

In their survey of emotion detection methodologies, Canales & Martínez-Barco outline four "bipolar sets" of emotions that are typically used in emotion categorization approaches:

"joy vs. sadness, anger vs. fear, trust vs. disgust, and surprise vs. anticipation" [116]. WordNet-Affect [117] and the NRC lexicon [118] are representative examples of lexicons that can be used to classify word polarity and emotion. Yadollahi et al.'s review of these and additional lexicons used in emotion mining shows that methodologies vary widely, reflecting the complex nature of human emotion which makes quantification a difficult task [115].

Of particular note in this research is the problem of using modern lexicons to study historic text. As an example, the NRC lexicon was developed in 2010 using Amazon Mechanical Turk, which means that emotions are labeled according to modern usage [118]. In the case of historic text, word usage can vary across time and culture [119]. Another issue is the original language of the text in question. While there exist lexicons developed for languages other than English, such as the FEEL lexicon for French [120], there remains significant work in order to develop a wider range of language choices [121]. When better options are unavailable, Araujo et al. find that use of English language sentiment analysis methodologies on machine translated text is often better than using less robust approaches on the original language [122].

Finally, given the overarching theme of identity-based conflict and genocide in this research, one cannot ignore the importance of hate speech and ethnic slurs. A number of studies have worked to develop new methods of detecting and classifying hate speech in text [123, 124, 125]. Natural language processing technologies and methodologies in this area continue to advance; however, one can expect the work will never reach an end point given the complexity and continual evolution of human language.

5.1.2 Why Sentiment Analysis for Eris?

In the Eris model, an agent's Ecosystem and Egosystem attributes can change based on the influence of its neighbors. The difference between these attributes then determines the probability that the agent will choose to be an active, passive, or complicit bystander (see Algorithm 2). This probability can be influenced in one direction or another depending on micro- and macro-level factors of escalation and restraint as defined in Sections 4.2 and 4.3.

The critical question here is which source of data can most appropriately inform model parameters in the validation cases? While Figure 2.2 shows that factors such as a strong economy can restrain genocide, this does not connect well with the the motivations and affective states of agents in the model as outlined in Table 2.2. Alternatively, the data contained in presidential speeches from Rwanda and Côte d'Ivoire are more suitable for the Eris model [12]. Through the use of sentiment analysis, these speeches can provide a longitudinal measure of emotion from an elite source representing the in-group. This naturally aligns with the theoretical framework presented in Table 2.2, given an assumption that emotion expressed on the part of societal elite has a level of influence on civilians [7]. This leads to the following hypotheses which inform the methodology for analysis leading toward model validation as outlined below in Section 5.1.3:

Hypothesis 1: **Relevant emotions can be measured in elite speech and used to provide a new measure of global factors of restraint in a society that affect all in-group agents.** In the 21st century, it is possible to obtain vast quantities of micro-level language data; however, when studying historic cases prior to the digital age, the text available is far less granular. Returning to the definition of β in Section 4.2.2, we will use measures of these emotions to inform the model's global β function representing factors of restraint in a society as understood by Straus [7, 8]. If the approach of sentiment analysis on Straus' set of presidential speeches is successful, the resulting index of how the relevant emotions change in time across cases can be applied to the model as a relative measure to inform β, factors of restraint.

Hypothesis 2: **Comparing the sentiment in the speech of these two presidents, and other elite actors during their terms, can provide a quantitative, relative measure of global factors of restraint in each respective society.** Each president had a long-term opportunity to convey information, thoughts, and ideas to his people. Straus finds the influence of the elite to be an important causal factor in genocide [7], and the goal of this aspect of our research is to use his database of speeches to measure emotions in the content as they relate to the ecosystem component of the Eris model (see Sections 2.3.2, 3.5, and 4.3.1).

5.1.3 Methodology for Measuring Ecosystem Motivation in Speeches

In order to develop a relative measure of global factors of restraint, we have elected to focus on how ecosystem orientation is expressed in elite speech. At this time, we have performed all sentiment analysis on English text only, translating from French when required. The NRC lexicon provides associations of the following emotions for English words: anger, fear, anticipation, trust, surprise, sadness, joy, and disgust [118, 126]. From this set of emotions, we hypothesize that ecosystem orientation can be quantified by measuring the level of joy and sadness in the *absence* of anger and fear in a given section of text. This leads to the following general equation:

$$ecosystem = [(joy + sadness) - (anger + fear)]/totalWords \tag{5.1}$$

where joy, sadness, anger, and fear are the sum of words expressing each of those emotions in the text.

The following is the methodology applied to the speeches in order to extract the ecosystem measure expressed in Equation 5.1:

1. Digitize each speech document from the database and convert to plain text.
2. Discard all data that is not from presidential discourse or other elite representatives. For interviews, this includes removing interview questions and only retaining the response.
3. When appropriate, split files that contain speech content from multiple events so that each is represented individually.
4. As necessary, translate files to English with Google Translate [127].
5. Using the NRC lexicon [126] on the English files, count the number of words classified as conveying joy, sadness, anger and fear for each paragraph in a given document.
6. Calculate the ecosystem value for that paragraph using Equation 5.1.

7. Average the paragraph level scores to determine the document level "Ecosystem Mean" score.
8. Plot the final scores for both Rwanda and Côte d'Ivoire to determine patterns and relative differences that can be applied in the Eris model.

Limitations. Section 5.3 presents the results of using the above methodology to analyze the selected speeches. While this work does yield quantitative measures that can be useful in validating the Eris model, it is essential to recognize limitations of the approach. These are as follows:

- The analysis is done only on English text. French files were machine translated, and this process results in some loss of meaning and accuracy.
- The NRC lexicon classifies associated emotions according to modern language usage, which is not necessarily accurate when applied to text from Sub-Saharan Africa from decades extending back to the 1960s.
- The lexicon does not account for the use of hate speech in these cases, particularly the appearance of the word "inyenzi" in the Rwandan data, which is a derogatory term commonly translated as "cockroaches," and used to dehumanize Tutsis from the 1960s [128, pp. 114-115]. Analysis of hate speech is a research area in itself, as discussed above in Section 5.1.1.
- These relative scores do not currently have a point of comparison from the larger historical record. Comparative emotional analysis of speeches from figures like Hitler and Churchill, and others, may yield essential, baseline measures.

5.2 African Presidential Speech Data

In developing his theory of the dynamics of genocide, Scott Straus collected and made available presidential speech data from various Sub-Saharan African nations [12]. These data represent elite speech over extended time periods in countries that experienced genocide, or did *not* experience one when it was expected. Straus analyzes these speeches in his book "Making and Unmaking Nations: War, Leadership, and Genocide in Modern Africa" [8], examining "themes" prior to conflict that may or may not have reappeared during crisis periods. These speeches are from a variety of sources including newspapers, summaries of speeches from foreign news services, and presidential interviews [12].

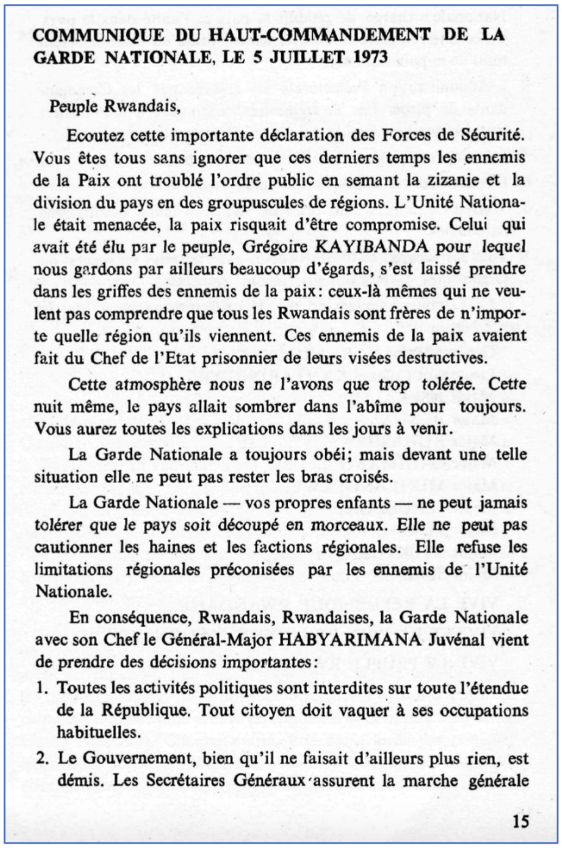
Straus provides speeches for six Sub-Saharan African countries: Côte d'Ivoire, Mali, Rwanda, Senegal, and Sudan. Here, only Rwanda and Sudan qualify as positive cases of genocide, while the others are negative [12]. Through careful exploration and comparative analysis of these documents, Straus theorizes that critical variables, often developed long before an outbreak of violence, determine the likelihood of the emergence of genocide [8].

Straus' research and analysis of these data inherently required the physical reading of historic documents by a subject-matter expert [7, 8]. As a supplement to the invaluable work of Straus and other historians and social scientists in interpreting and contextualizing historic texts, we seek to use computational techniques to provide additional, quantitative information. While the results of the computational analysis are used here for the purpose of informing the Eris model, we hope that the digitized version of a subset of these speeches and analytical approach will prove useful to historians and social scientists.

5.2.1 Overview of Speech Data

Figure 5.1 below shows two representative examples from the database of presidential speeches [12]. The database includes speeches with both greater and lesser resolution and clarity than these examples; however, the layout and quality differences between these two documents sufficiently conveys the range of difficulty we anticipated in working to digitize these documents.

Rwanda speech overview. Straus' database of Rwandan presidential speeches extends from October 26, 1960 to February 26, 1994 [129]. We have chosen to analyze only speeches from President Juvénal Habyarimana and others acting as his representatives. Habyarimana became president of Rwanda as the result of a coup in 1973 [130], and the first of his speeches in the database is from July 5th of that year. Of the 63 documents analyzed here, all were speeches or interviews from President Habyarimana except for one speech from



COMMUNIQUE DU HAUT-COMMANDEMENT DE LA GARDE NATIONALE, LE 5 JUILLET 1973

Peuple Rwandais,

Ecoutez cette importante déclaration des Forces de Sécurité. Vous êtes tous sans ignorer que ces derniers temps les ennemis de la Paix ont troublé l'ordre public en semant la zizanie et la division du pays en des groupuscules de régions. L'Unité Nationale était menacée, la paix risquait d'être compromise. Celui qui avait été élu par le peuple, Grégoire KAYIBANDA pour lequel nous gardons par ailleurs beaucoup d'égards, s'est laissé prendre dans les griffes des ennemis de la paix: ceux-là mêmes qui ne veulent pas comprendre que tous les Rwandais sont frères de n'importe quelle région qu'ils viennent. Ces ennemis de la paix avaient fait du Chef de l'Etat prisonnier de leurs visées destructives.

Cette atmosphère nous ne l'avons que trop tolérée. Cette nuit même, le pays allait sombrer dans l'abîme pour toujours. Vous aurez toutes les explications dans les jours à venir.

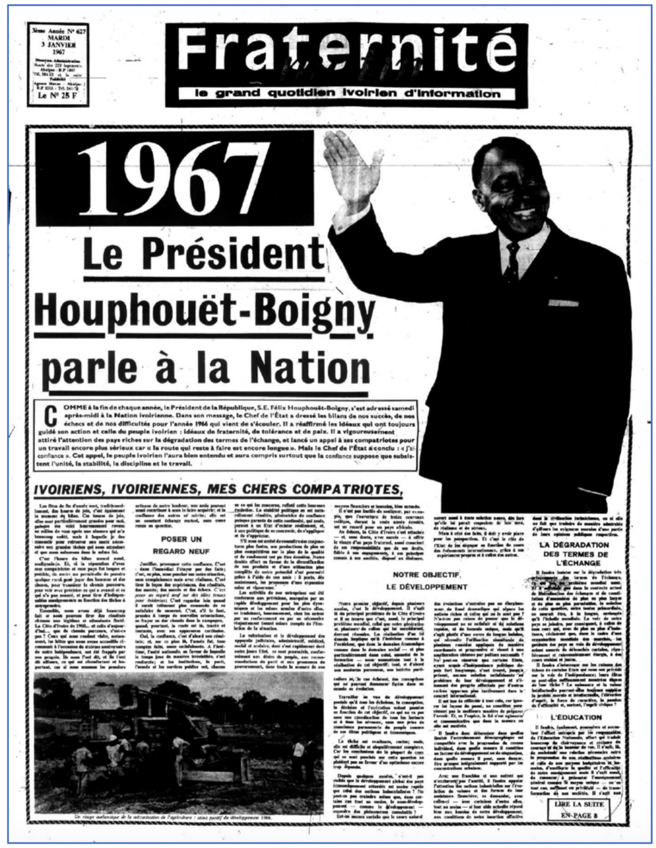
La Garde Nationale a toujours obéi; mais devant une telle situation elle ne peut pas rester les bras croisés.

La Garde Nationale — vos propres enfants — ne peut jamais tolérer que le pays soit découpé en morceaux. Elle ne peut pas cautionner les haines et les factions régionales. Elle refuse les limitations régionales préconisées par les ennemis de l'Unité Nationale.

En conséquence, Rwandais, Rwandaises, la Garde Nationale avec son Chef le Général-Major HABYARIMANA Juvénal vient de prendre des décisions importantes:

1. Toutes les activités politiques sont interdites sur toute l'étendue de la République. Tout citoyen doit vaquer à ses occupations habituelles.
2. Le Gouvernement, bien qu'il ne faisait d'ailleurs plus rien, est démis. Les Secrétaires Généraux assurent la marche générale

15



3ème Année N° 627
MARDI
3 JANVIER
1967

Le N° 25 F

Fraternité

le grand quotidien ivoirien d'information

1967

Le Président Houphouët-Boigny parle à la Nation

COMME à la fin de chaque année, le Président de la République, S.E. Félix Houphouët-Boigny, s'est adressé samedi après-midi à la Nation ivoirienne. Dans son message, le Chef de l'État a dressé les bilans de nos succès, de nos échecs et de nos difficultés pour l'année 1966 qui vient de s'écouler. Il a réaffirmé les idéaux qui ont toujours guidé son action et celle du peuple ivoirien : idéaux de fraternité, de tolérance et de paix. Il a vigoureusement attiré l'attention des pays riches sur la dégradation des termes de l'échange, et lancé un appel à ses compatriotes pour un travail encore plus sérieux car « la route qui reste à faire est encore longue ». Mais le Chef de l'État a conclu : « J'ai confiance ». Cet appel, le peuple ivoirien l'aura bien entendu et aura compris surtout que la confiance suppose que subsistent l'unité, la stabilité, la discipline et le travail.

IVOIRIENS, IVOIRIENNES, MES CHERS COMPATRIOTES,

POSER UN REGARD NEUF

NOTRE OBJECTIF, LE DÉVELOPPEMENT

LA DÉGRADATION DES TERMES DE L'ÉCHANGE

L'ÉDUCATION

LIRE LA SUITE EN PAGE 8

Figure 5.1: Samples of documents in Straus' "African Presidential Speeches Database" [12]. Left: Rwanda, 5-Jul-1973, Right: Côte d'Ivoire, 3-Jan-1967.

Prime Minister Nsanzimana (11-Mar-1992), and one from Leon Mugesera (22-Nov-1992). Leon Mugesera was an advisor to Habyarimana's political party, the National Republican Movement for Development and Democracy (MRND) [131], and this speech is considered to be the "rally call" for the Tutsi geocide [132].

Côte d'Ivoire speech overview. Félix Houphouët-Boigney was Côte d'Ivoire's first post-colonial president. He took office in 1960, and remained president until his death in 1993 [133]. Straus' database for Côte d'Ivoire begins with Houphouët-Boigney's speech on August 6, 1961, and includes speeches, reports and interviews after his death, through the transitional period, ending with a foreign news report on the status of an early election under President Laurent Gbagbo on October 30, 2005 [134]. As explained in the comparative analysis section below, we have only included speeches from Côte d'Ivoire prior to Houphouët-

Boigney's death. The last speech in this subset is from October 6, 1990. Of the 46 documents analyzed here, all are from President Houphouët-Boigney except for two speeches from the President of the National Assembly, Phillip Yacé (6-Aug-1966, 2-May-1969), two from a later President of the National Assembly, Henri Konan Bédié (1-May-1987, 27-Apr-1988), and one from Minister of State, Mathieu Ekra (1-May 1988).

Notes on comparative analysis. As detailed above, we elected to focus only speeches from Presidents Habyarimana and Houphouët-Boigney, or other elites in their administrations who were included in the database. We chose *not* include analysis of the speeches Straus provides from Rwanda's Gregoire Kayibanda (October 26, 1960 - January 1, 1973), or from other political figures in Côte d'Ivoire after the death of Houphouët-Boigney [129]. This filtering of the data allows for increased consistency in the comparison between Rwanda and Côte d'Ivoire. Habyarimana's presidency extended over approximately two decades, with the 1994 genocide occurring shortly after his assassination. Houphouët-Boigney's presidency was over three decades long, with civil tensions arising in the decade after his death.

5.2.2 Digitization of Documents and Descriptive Statistics

Conversion of documents to plain text. The vast majority of the speeches in Straus' database are scans of newspaper articles and other documents. For documents of sufficiently high quality, we used AABBY Fine Reader to perform optical character recognition and convert the file to plain text [135]. Even for the highest quality original, there was still a chance of error in the conversion process. As such, we performed manual review and corrections of all digitally converted files. For documents that could not be successfully read and converted by the software, we performed manual transcription. For the Rwanda data, only four of 63 documents required manual transcription. In contrast, 36 of the 46 speeches from Côte d'Ivoire had to be manually transcribed due to very poor resolution in the scans.

Descriptive statistics. Table 5.1 below presents the date ranges and descriptive statistics for the speech data extracted from Straus' database. Recall from the methodology outlined in Section 5.1.3 that documents were cropped as necessary in order to produce a data set the represents, as closely as possible, *only* the speech of the societal elite. As such, the number of files included and their length may differ from the contents of the original database.

Table 5.1: Descriptive statistics for analyzed speeches from Rwanda and Côte d'Ivoire. All documents are taken from Scott Straus' "African Presidential Speech Database" [12].

Descriptive Statistic	**Rwanda** *5-Jul-1973 to 26-Feb-1994*	**Côte d'Ivoire** *6-Aug-1961 to 6-Oct-1990*
Number of Documents	63	46
Minimum Word Count	224	114
Maximum Word Count	13798	4772
Mean Word Count	3106	1629
Standard Deviation	3471	1330

This concludes the overview of the speeches selected for this analysis. Next, we present the results of the sentiment analysis for each case in Sections 5.3.1 and 5.3.2. These results work toward establishing the validity of Hypothesis 1 posed in Section 5.1.2 by concretely identifying a new, quantitative measure of elite emotion that can inform the level of restraint in a society against out-group persecution. Section 5.3.3 compares these results, working toward an answer to Hypothesis 2 by yielding a relative measure of these composite scores across societies. This result can be taken as an index that can be used to inform the parameters of the Eris model for individual cases as explained in Section 5.3.4.

5.3 Results From Sentiment Analysis

Recall from the methodology outlined in Section 5.1.3 that we used the NRC lexicon to compute paragraph level ecosystem scores as defined in Equation 5.1. In the results detailed below, each data point represents the average of those paragraph level scores across the document. The following three subsections cover the following: results for Rwanda, results for Côte d'Ivoire, and a comparative analysis of the results for both cases.

5.3.1 Rwanda Results

We begin with an examination of the individual composite scores used in Equation 5.1 to calculate an ecosystem mean score for every document in the data set. Figure 5.2 below shows these results for Rwanda. Of note here is that $(joy + sadness)$ does, on average, remain higher than $(anger + fear)$, and larger gaps between the two indicate that emotion in the speech is weighted more heavily in one direction than the other.

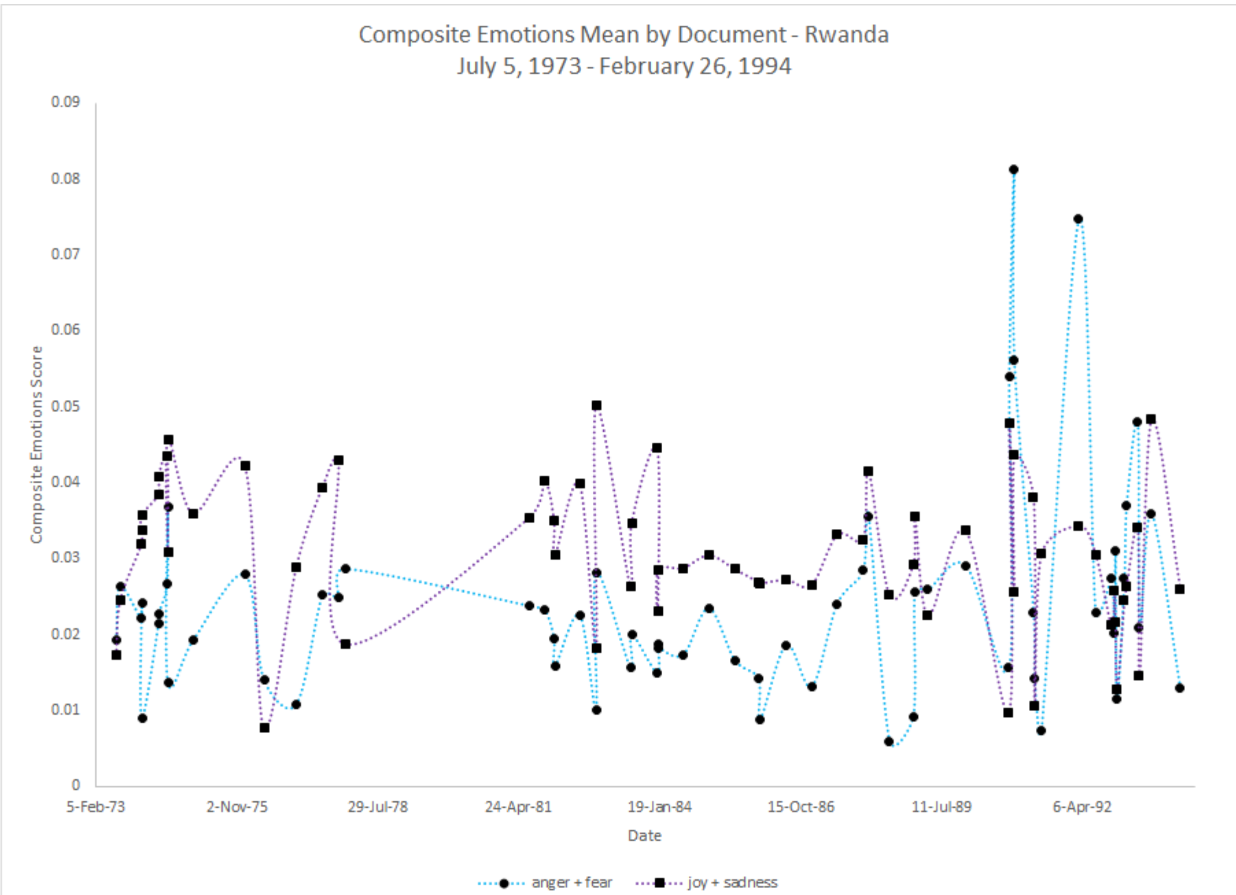


Figure 5.2: Composite score means by document for Rwanda: $(anger + fear)$ compared to $(joy + sadness)$.

Figure 5.3 below shows the raw ecosystem score by date, computed with Equation 5.1, for each document in the Rwandan data set. The overall trend for all data, shown by the regression line, is moderately negative; however, it contains two distinct sub-trends. There is a relatively constant, positive trend after the ratification of the new constitution and election of Habyarimana in 1978. This trend changes dramatically both with respect to the polarity and volatility after rebel forces invaded Rwanda from Uganda starting in 1990. It is important to note that a zero ecosystem score is a critical threshold. An ecosystem score falling below zero shows emotional content that has an overall weighting toward fear and anger, negating the effects of joy and sadness.

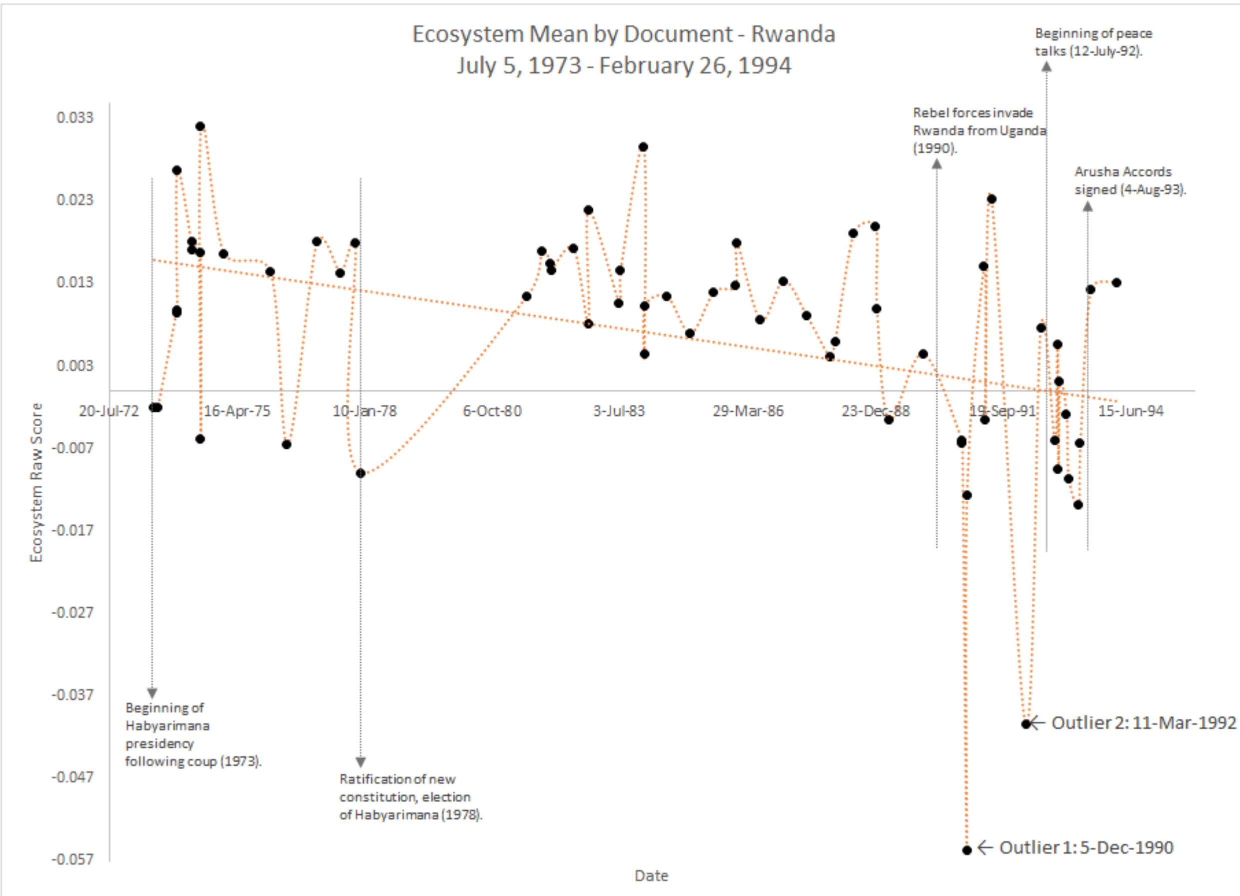


Figure 5.3: Ecosystem mean, raw score by document for Rwandan speeches with corresponding outliers and events of interest.

There are two outliers to consider in these data: Outliers 1 and 2 marked in the chart above. Below is an overview of the content and significance of these speeches, followed by a quantification of their effect on the results.

Outlier 1: 5-Dec-1990, 1005 words [136]. This document is a BBC translation of a Radio Rwanda broadcast summarizing President Habyarimana's comments to the Central

Committee of the MRND about the ongoing war. He explains steps the government is taking to successfully defeat the rebels. While the report contains few direct quotes, it comes from the only radio station in Rwanda at the time. Radio Rwanda was considered "very much the voice of the government and of the president himself" until 1992 [27, p. 67], which justifies this broadcast being considered elite communication in the context of this research. An excerpt from this BBC translated report is: "Since that defeat, the enemy has now adopted the tactics of invading secretly, killing the people and destroying things and then go back from where he came from" [136]. This is a prime example of language expressing fear.

Outlier 2: 11-Mar-1992, 489 words [137]. This document is an FBIS translation of Prime Minister Nsanzimana's speech in which he outlined seven "appropriate measures" the government would be taking in order to address the "ongoing war at the northern border of our country and the troubles in the commune of Kanzenze." While the directives in this speech are intended to calm the violence, they are posed in a manner that expresses a high degree of fear and anger, such as: "Everyone must refrain from any act which might violate human rights, and those guilty of such acts must be severely punished" [137].

While both of these files are relatively small compared to the mean word count for the Rwandan data (see Table 5.1), each is clearly applicable to the topic the elite communication with citizens. Figure 5.4 and Table 5.2 below show the effect of these outliers on the overall results for the data set.

Table 5.2: Analysis of Outliers 1 and 2 effect on Rwanda speech statistics.

	Mean	Standard Deviation	Regression Slope	Regression Intercept
All Data	0.0067	0.0147	-2.2678 E-06	0.0159
Excluding Outliers	0.0084	0.0109	-1.5972 E-06	0.0149
Percent Change	25%	-26%	30%	-6%

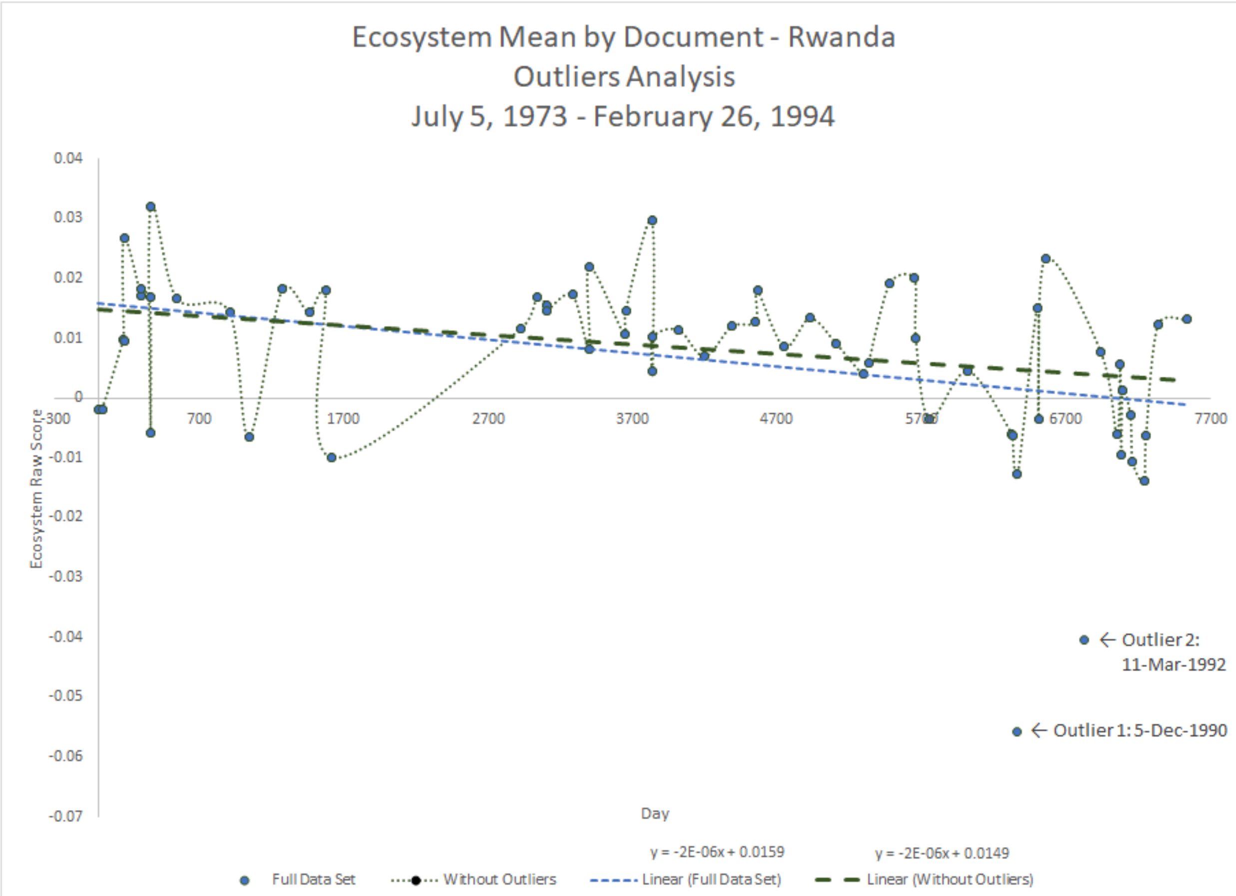


Figure 5.4: Regression lines with and without Outliers 1 and 2 in Rwanda data.

It is clear that these outliers do have a significant effect on the overall result, causing the mean to be 25% higher, and the regression line slope to be 30% less negative. This shows that these outliers cause a lower mean value and a steeper negative slope, predicting a greater decrease in ecosystem value over time than when they are excluded. However, as explained above, their content is fully representative of elite communication to its citizenry. As such, we have elected to retain these outliers in the validation process outlined in Section 5.3.4.

5.3.2 Côte d'Ivoire Results

Next, we present an overview of the results from sentiment analysis of the Côte d'Ivoire data. First are the composite scores for $(anger + fear)$ and $(joy + sadness)$ shown in Figure 5.5 below. Here $(joy + sadness)$ remains, on average, higher than $(anger + fear)$, as was the case in the Rwandan data. However, in this case there are fewer documents in which $(anger + fear)$ is greater than $(joy + sadness)$, and there appear to be sections with larger gaps between these two composite emotions. Looking ahead, Section 5.3.3 more closely examines and quantifies the differences between the Rwandan and Ivorian results.

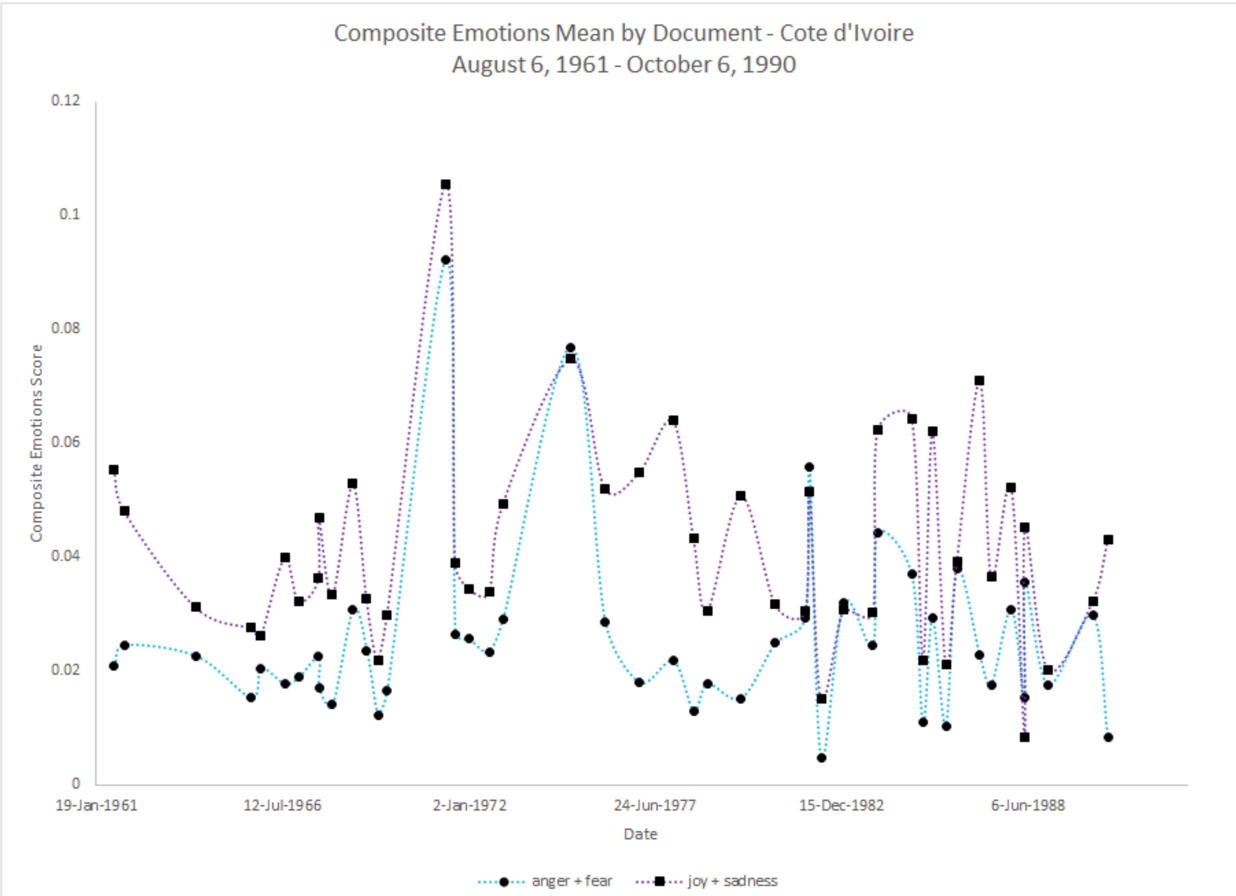


Figure 5.5: Composite score means by document for Côte d'Ivoire: $(anger + fear)$ compared to $(joy + sadness)$.

Figure 5.6 below shows the raw ecosystem scores for each document in the Côte d'Ivoire data set by date. While volatile, the overall trend is only slightly negative, with the vast majority of documents returning a positive value for ecosystem. There are only four documents with an ecosystem score than falls below zero, indicating little tendency for elite speech to drift across the neutral value toward $(anger + fear)$ weighting.

Again there are two outliers to consider in these data: Outliers 1 and 2 marked in the

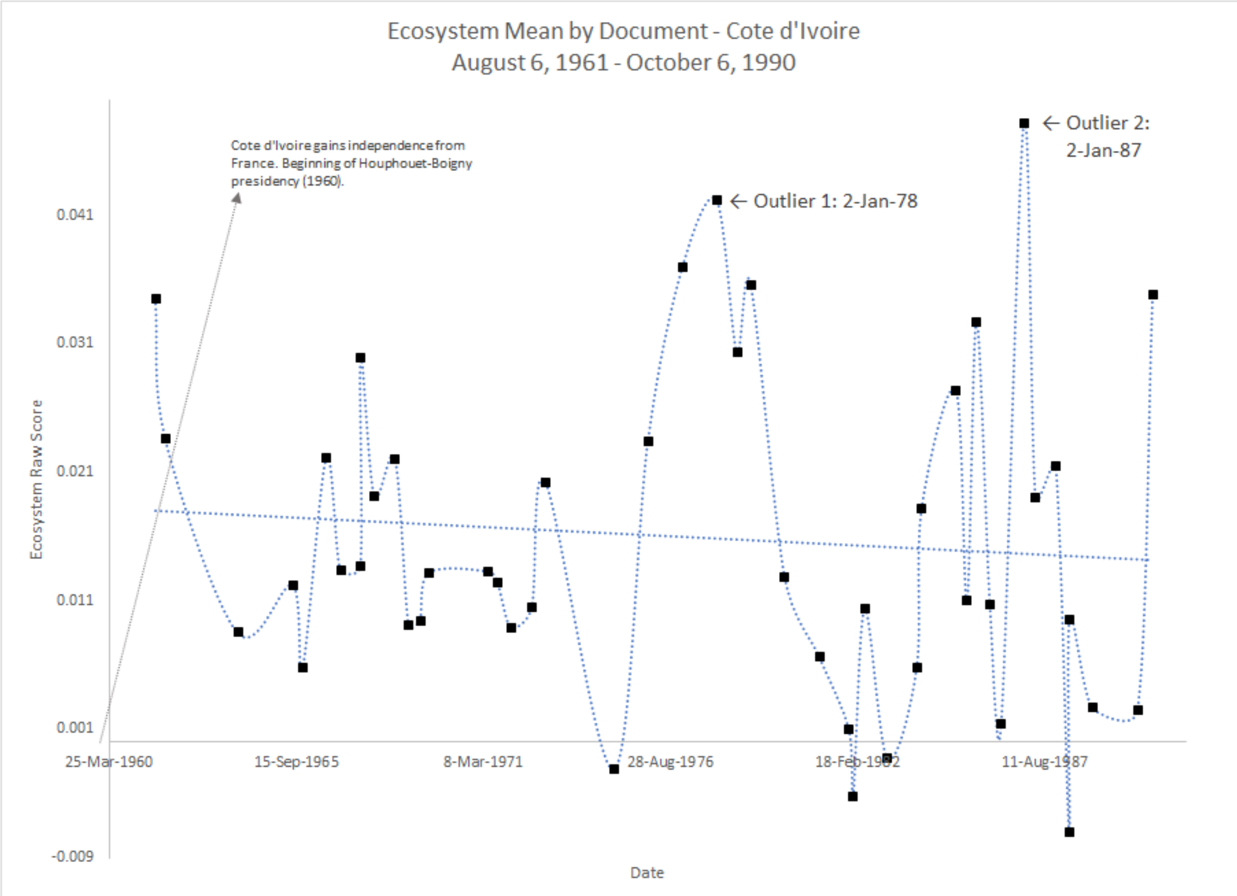


Figure 5.6: Ecosystem mean, raw score by document for Côte d'Ivoire speeches with corresponding events of interest.

chart. Below is an overview of the content and significance of these speeches, followed by a quantification of their effect on the results.

Outlier 1: 2-Jan-1978, 455 words [138]. This is President Houphouët-Boigney's New Year's speech. There are 19 speeches in the Côte d'Ivoire database that were given at the start of the new year, representing about 40% of the total data set. As such, this speech is highly representative of what is contained in the data as a whole. A machine translated excerpt from this speech provides an excellent example of why Houphouët-Boigney's dialogue measures so high in ecosystem: "Continue down this path, for ourselves and for others. For ourselves, because our sacred duty is to prosper Ivorian homeland and ensure the future for others, because I believe in the virtue of the example and we need to bring in a massive way and our selfless help our unfortunate brethren" [138].

Outlier 2: 2-Jan-1987, 354 words [139]. Again, this is a speech given at the beginning of the new year, and is an excellent candidate for remaining included in the data despite its ecosystem value being the highest of all documents. Another machine translated excerpt helps explain this high ecosystem result: "It is in peace, harmony and the work we will continue our work, that God's help will always make more fruitful if we see the virtues of charity and justice to help the afflicted to endure their condition and justify success those that fate has favored when they know share the benefits with more unhappy" [139].

As was the case with the Rwandan outliers, these speeches are relatively small compared to the mean word count for the Ivorian data (see Table 5.1). However, each remains an appropriate example of elite communication to the nation's citizens, and are also well placed with other speeches given at the start of the new year. Figure 5.7 and Table 5.3 below show the effect of these outliers on the overall results for the data set.

Table 5.3: Analysis of Outliers 1 and 2 effect on Côte d'Ivoire speech statistics.

	Mean	Standard Deviation	Regression Slope	Regression Intercept
All Data	0.0159	0.0128	-3.6048 E-07	0.0180
Excluding Outliers	0.0146	0.0114	-6.5817 E-07	0.0183
Percent Change	-8%	-11%	83%	2%

These outliers have a significant effect on the overall result, particularly with respect to the change in slope of the regression line. The slope of the line when outliers are excluded becomes 83% more negative. This is an even greater effect than the absolute change in the slope for the Rwandan data. It is clear that these two outliers cause the regression line slope to be significantly more positive when they are included in the analysis, predicting higher levels of ecosystem values over time than when they are excluded. Nevertheless, these

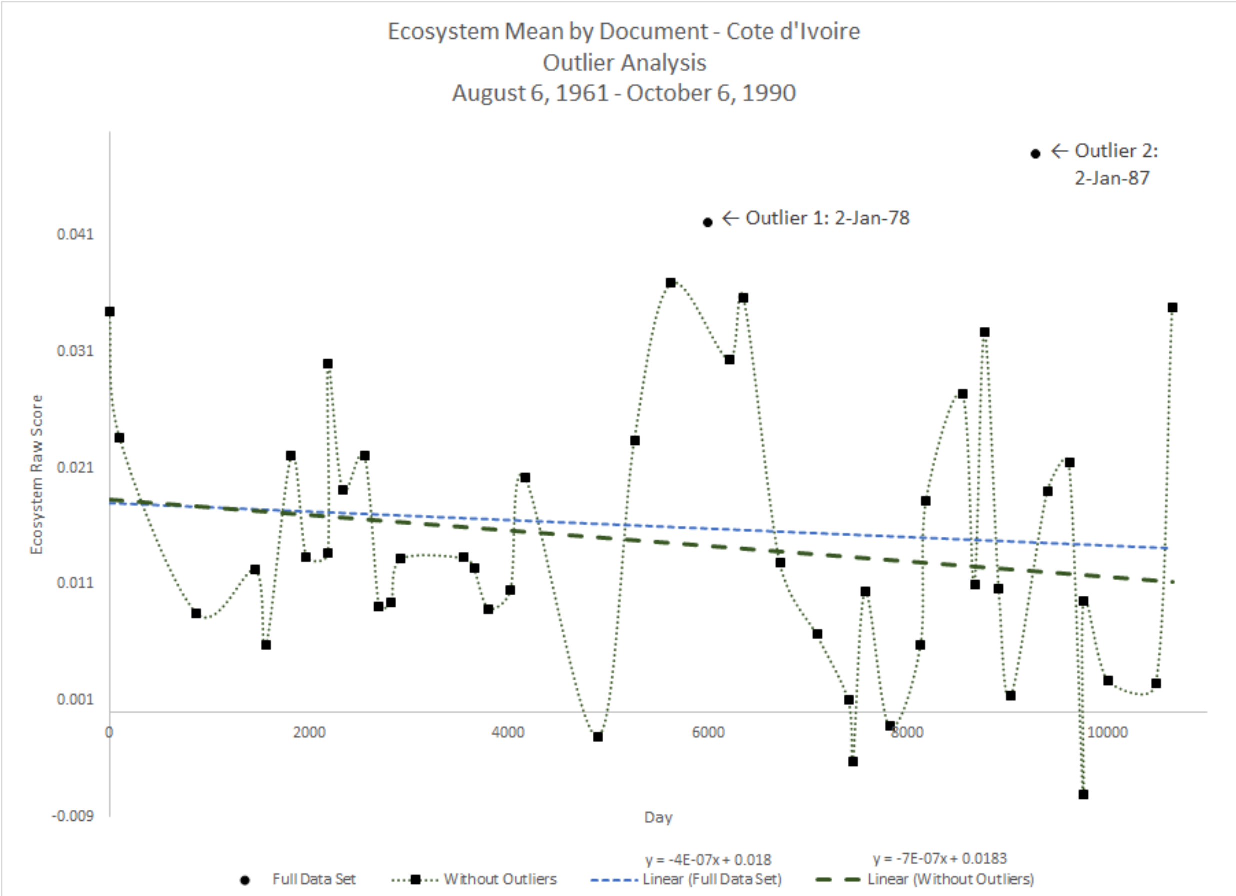


Figure 5.7: Regression lines with and without Outliers 1 and 2 in Côte d'Ivoire data.

speeches are entirely appropriate for the data set as a whole, as explained above. We have again elected to retain these outliers in the validation process detailed in Section 5.3.4.

Given the above, the next step is to determine the relative meaning of these results and analyses. The purpose of doing so is to quantify the differences between these two cases in a way that can inform the β function of the Eris model, the global factors of restraint in the environment, for the purpose of empirical validation.

5.3.3 Comparison of Results Across Cases

The first step in comparing the results for these two cases involves determining whether or not the data support the hypothesis that Equation 5.1 will yield a meaningful measure of ecosystem motivation. Figure 5.8 below compares how the emotional scores of $(anger + fear)$ and $(joy + sadness)$ taken individually differ between the two cases. Note that the x-axis is transformed from the dates shown in Figures 5.3 and 5.6 to absolute days. The first speech in each series is Day 1, with the following speech plotted according to how many days are between it and the preceding speech. Note that the Ivorian data, while fewer speeches in total, extends further in time than those from Rwanda. Houphouët-Boigney's presidency lasted approximately ten years longer than that of Habyarimana, yielding this longer timeline.

Figure 5.8: Comparison of $(anger + fear)$ and $(joy + sadness)$ across cases.

Starting with the $(anger + fear)$ comparison chart on the left, we can see that both sets of results show similar levels of these emotions, with the Rwandan data having a greater number of outliers overall. Alternatively, the $(joy + sadness)$ comparison chart shows that the Ivorian data typically has higher values, with the Rwandan data showing no extreme outliers. This preliminary view of these results shows that $(anger + fear)$ alone is insufficient to distinguish between these two cases. Both show nearly equivalent levels of these two emotions, which does not correspond to what Straus highlights as distinct differences between the communications of these presidents [7, 8]. The $(joy + sadness)$ results align more closely to Straus' findings; however, there is not a significant difference in trends other than to say that the results are generally higher for the Ivorian case.

We next examine the results through the lens of Equation 5.1, where ecosystem motivational orientation is measured as $(joy + sadness)$ in the *absence* of $(anger + fear)$. Figure

5.9 shows the differences between cases using this measure. In the figure, we can see that the regression line for Côte d'Ivoire has an almost constant slope, while that for Rwanda is significantly more negative. This means that over time, the emotion expressed in the elite speech from Rwanda weighs more strongly toward anger and fear than was the case in Côte d'Ivoire. We can also see that outliers from Côte d'Ivoire that measured high in $(anger + fear)$ from Figure 5.8 are moderated by the presence of $(joy + sadness)$ in those same speeches. These results more clearly align with Straus' theories regarding the tone of communication from these leaders [7, 8].

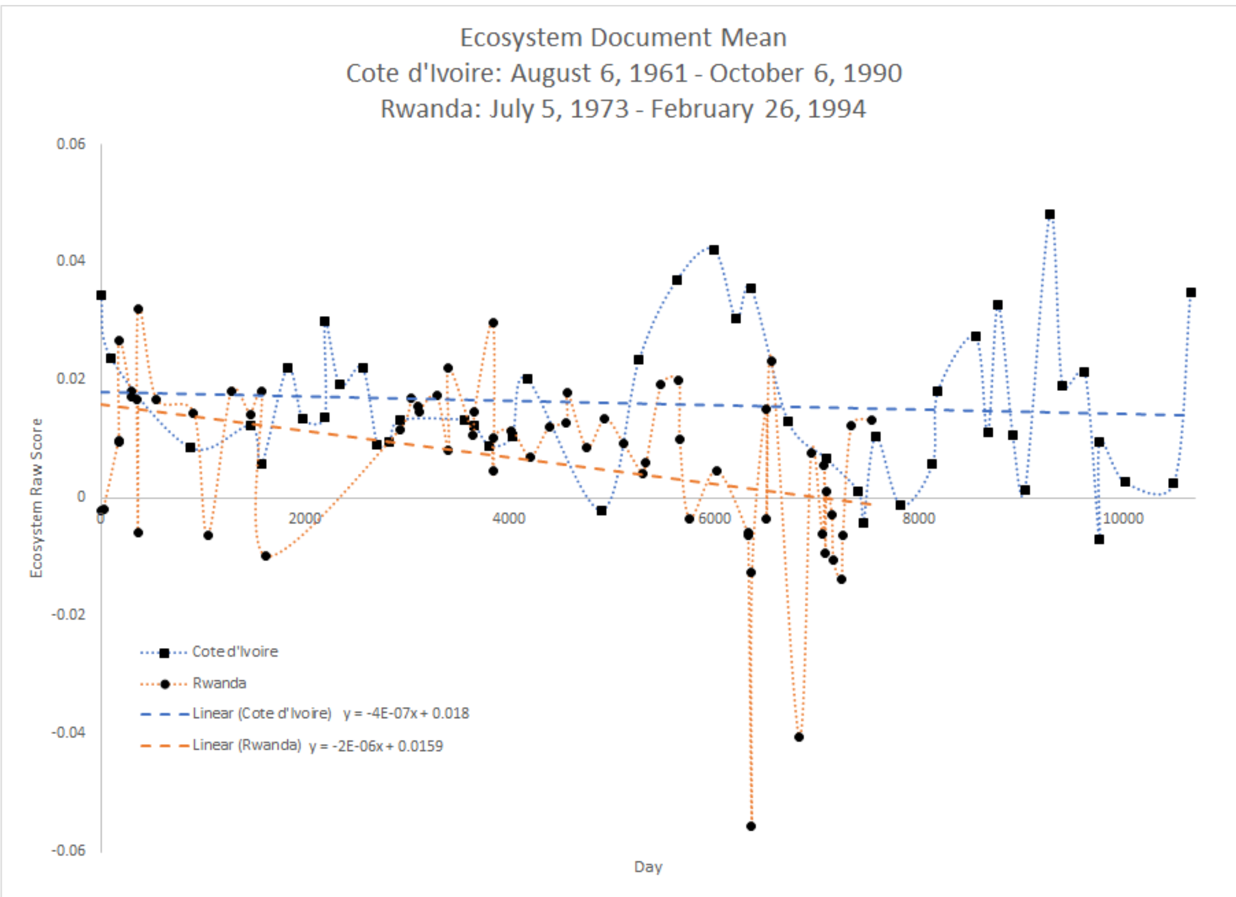


Figure 5.9: Combined ecosystem mean, raw scores for Rwanda and Côte d'Ivoire.

Table 5.4 quantifies the differences between ecosystem results from each data set. This clarifies the magnitude of differences between these two data sets. The mean ecosystem value for Côte d'Ivoire is 141% higher than that for Rwanda, and the slope of its regression line, while still negative, is 84% *less* steep. Included in this table is the exact number of speeches from each case with an ecosystem value below zero, which is a critical threshold. A document with $ecosystem = 0$ means that the emotional content of that speech is effectively neutral with respect to this measure. A value below zero means that $(anger+fear)$ outweigh $(joy+sadness)$. In the case of the Rwandan results shown in Figure 5.3, 13 of the 18 negative ecosystem scores occur *after* the start invasions from Uganda in 1990. The results confirm what we would expect to see in a country experiencing such incursions, as they are a threat to the elite power structure (see Figure 2.2), which causes self-preservation motivations to arise (see Section 2.3.2 and Table 2.2).

Table 5.4: Comparison of descriptive statistics and regression analysis between cases.

	Mean	Standard Deviation	Regression Slope	Regression Intercept	Number of Docs with Ecosystem <0	Percent of Docs with Ecosystem <0
Rwanda	0.0066	0.0147	-2.2678 E-06	0.0159	18 of 63	29%
Cote d'Ivoire	0.0159	0.0128	-3.6048 E-07	0.0180	4 of 46	9%
Percent Change (Rw to Cote)	141%	-13%	-84%	13%		

Next, we look more closely at a comparison of the mean results. While analysis of individual ecosystem scores and their regression lines is a more nuanced approach, a simple examination of differences in means can reveal important differences. Given the broader goal of using these results to validate the Eris model, it is useful to begin with the most simple approach, increasing in complexity as necessary. To that end, Figure 5.10 below visualizes the magnitude of difference between the means from Table 5.4.

As we will discuss further in Section 5.3.4, it is possible to implement this simple differences between means in the Eris model in order to reproduce the Ivorian and Rwandan scenarios. However, given the high percentage of negative ecosystem scores and clear change in trend in the Rwandan scores after 1990 (see Table 5.4 and Figure 5.3), it may be more accurate and true to the historic record to split these data. Figure 5.11 shows how doing so affects our

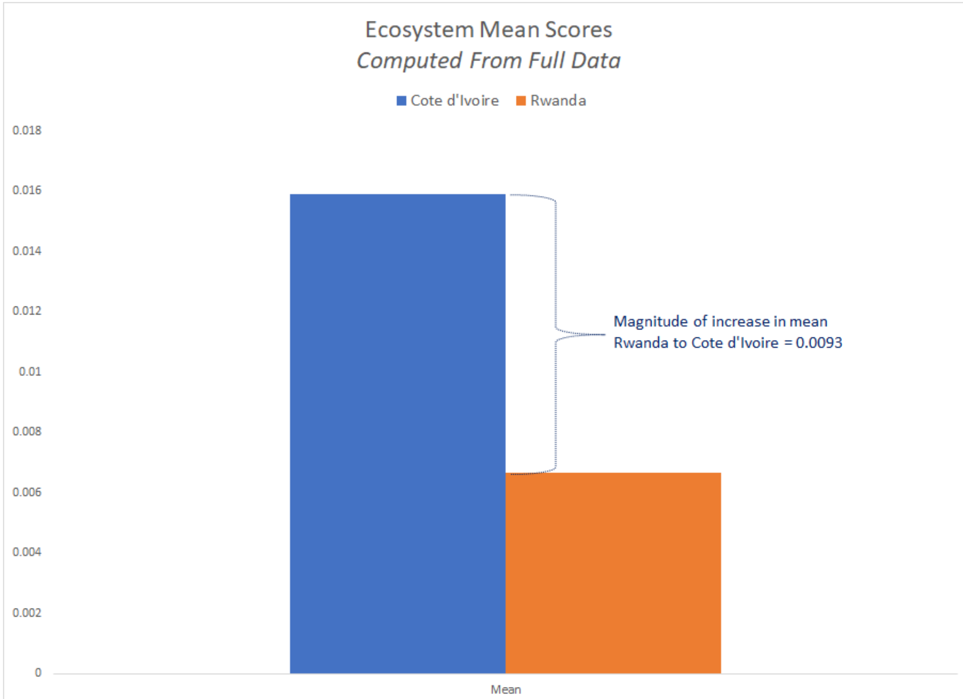


Figure 5.10: Ecosystem data set means for all data in each case.

understanding of the difference between these two cases. Here, the first mean for Rwanda is computed from all ecosystem document scores before 30-October-1990, and the second mean is computed from the document scores including and after that date.

Table 5.5 summarizes the data that inform Figures 5.10 and 5.11. These results provide the most simple quantification of differences between the Ivorian and Rwandan ecosystem scores. Returning to our two hypotheses stated in Section 5.1.2, the above results show that both were correct. First, the results presented for each individual case in Sections 5.3.1 and 5.3.2 show that we have been able to identify and measure "relevant emotions" in elite speech that can be used to inform the Eris model, confirming Hypothesis 1. We achieved this by focusing on a composite measure of emotion, ecosystem, as defined by Equation 5.1. Second, the comparison of results in Section 5.3.3 confirms the validity of Hypothesis 2. Comparison of the sentiment of the speeches across these two cases provides multiple options for relative measures of global level factors of restraint as expressed by elite ecosystem orientation. To close this chapter, we next present a simple framework for validating Eris model using the

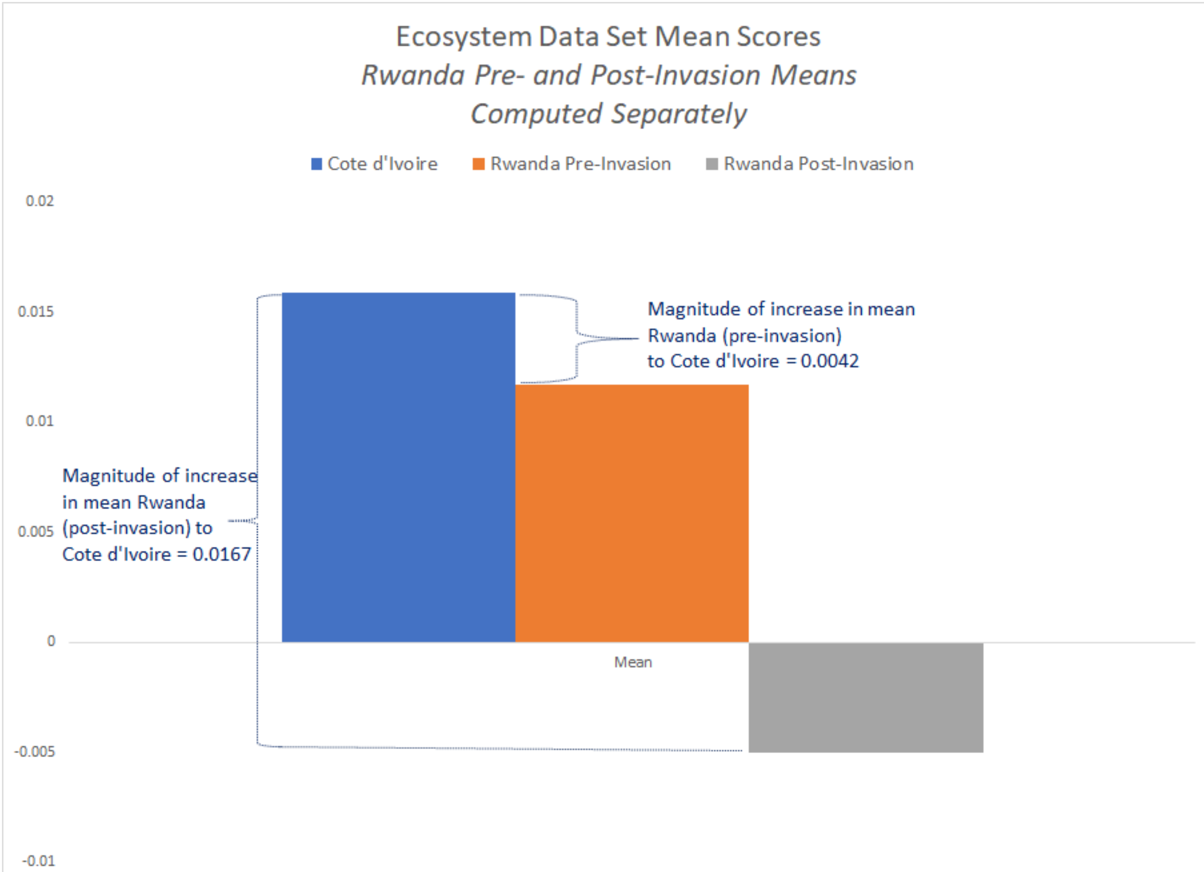


Figure 5.11: Ecosystem data set means for each case. Rwanda means divided after start of 1990 invasions from Uganda. Côte d'Ivoire means remain computed on all data.

above results.

5.3.4 Framework for Model Validation

As outlined in Section 4.2, the Eris model implements β as a global function representing societal factors restraining violence against, and persecution of, out-group populations. Having applied the methodology outlined in Section 5.1.3 to the speeches from Rwanda and Côte d'Ivoire found in Scott Straus' database [12], we now have appropriate measures of ecosystem emotion on the part of the societal elite for both cases.

Validation goals. Returning to the research gap stated in Section 3.4.3, general models are not scenario specific, and scenario specific models are not easily generalized, we seek a form of validation that neither exclusively empirical nor based solely on stylized facts. This

Table 5.5: Summary of ecosystem means comparison across cases. Includes results for Rwandan data split at start of 1990 invasions from Uganda.

	Dates	Number of Documents	Mean Ecosystem	Standard Deviation	Percent Change (Rw to Cote)
Cote d'Ivoire	6-Aug-1991 - 6-Oct-1990	46	0.0159	0.0128	-
Rwanda (all)	5-Jul-1973 - 26-Feb-1994	63	0.0066	0.0147	141%
Rwanda (pre-invasions)	5-Jul-1973 - 27-Dec-1989	44	0.0117	0.0090	36%
Rwanda (post-invasions)	30-Oct-1990 - 26-Feb-1994	19	-0.0050	0.0185	418%

falls in the "other forms of validation" category shown in Figure 3.1, and is defined by the following goals:

- The model should reproduce generalized patterns of identity-based conflict that are common to all scenarios.
- The model should *also* reproduce generalized patterns that differ between scenarios, particularly with respect to the emergence or non-emergence of genocide.

Validation framework. To achieve the above goals, we will calibrate the model for the Rwandan scenario, and then determine if the mean ecosystem values from the Ivorian data are sufficient to restrain genocide under those same circumstances. If this is the case, the modeling results will show that this level of ecosystem in the dialogue of the Ivorian elite is a potential explanation for the non-emergence of genocide in that case. The validation steps are as follows:

1. Initialize the model with a synthetic population that is representative of the overall population of the Rwanda in 1994.
2. Determine the β value that will allow genocide to emerge with a death rate and timeline similar to the Rwandan case.
3. Adjust the model's β value according to the Ivorian results to determine if the increase is sufficient to prevent a genocide.

4. Determine if there are conditions within the model environment that can negate any restraining effect of a revised β value.

If the above goals and simulation steps lead the Eris model to successfully simulate the death rate of the 1994 Rwandan genocide, and the non-emergence of genocide given an adjusted β value to reflect the Ivorian case, the model will have achieved a new form of validation. It can easily be made specific to a particular scenario, while remaining highly generalized and easily adaptable exploring different outcomes using β values informed by other cases.

Contributions. The following is a short summary of research contributions specific to the work detailed in this chapter:

1. Equation 5.1 defines ecosystem as a novel measure of composite emotions quantified through the use of sentiment analysis. The results of our analysis support the validity of this equation in determining an ecosystem motivational orientation in elite speech.
2. The methodology presented in Section 5.1.3, when applied to the selected cases and speeches from Straus' presidential database [12], yields a quantification of ecosystem that aligns with expected relative levels based on historic events.
3. The comparative results presented in Section 5.3.3 lead to a novel framework for validation of the Eris model through their application to its global β function.

The next two chapters detail two experimental sets performed with the Eris model described in Chapter 4. Chapter 6 presents results from experiments exploring generalized inter-group dynamics, including details regarding model verification. Chapter 7 presents validation experiments using the methodology outlined above. Both chapters aim to show the Eris model's usefulness for understanding both general and specific dynamics of identity-based conflict, including unveiling potential dynamics and restraints that *prevent* genocide from occurring.

CHAPTER 6: EXPERIMENTAL SET 1 – GENERALIZED INTER-GROUP DYNAMICS

The first of two experimental sets conducted with the Eris model explores generalized inter-group dynamics. These experiments are conducted without the use of the data and results explored in Chapter 5 in order to determine how the model behaves in a generic scenario. As explained in the model description in Chapter 4, this simplified, artificial society contains two identity groups: an in- and out-group. In-group agents interact and adapt with one another, accounting for different global conditions when determining their bystander choice. Perpetrators move randomly through the environment, and are a threat to out-group agent survival if there are not a sufficient number of active bystanders to offer protection. The results of this set of experiments show the capabilities of the Eris model to explain identity-based conflict in a very general sense. The environment is not calibrated to match any particular real-world event, as the purpose of these experiments is to determine if the model can reproduce general dynamics of the problem before beginning the process of empirical validation.

6.1 Methodology

While the details of each sub-experiment's methodology varies according to the research question and goals, they share a fundamental methodological framework. Each sub-experiment has the following in common:

- The model used is as described in Chapter 4 with the following exception: for the purpose of comparison of its effect, some experiments do not allow agent Susceptibility to moderate the impact of external factors on OrientationDifference in Algorithms 2, 3, and 4.
- Unless otherwise noted, results are the averages of 100 replicates. Section 4.6.3 details

the methodology used to determine the optimal number of replicates necessary for stable results.

- All experiments in this set use the following settings:
  - RadiusofSight = 10 NetLogo patches
  - ProbabilityofMutation = 1 in 10,000
  - ProbabilityofDeath = 1 in 25,000
  - SusceptibilityFraction = 10,000
  - MaximumDays = 15,000[1]
  - Environment size = 75x75 NetLogo patches

6.1.1 Overview of Results

The most significant results of this experimental set were as follows:

- Individual-level susceptibility to change according to global influences dramatically reduces the beneficial effect of factors of restraint (Section 6.3, Figure 6.5).
- As the level of factors of restraint in a society increase, they begin to have an exponentially beneficial impact on protection of the out-group from persecution (Section 6.3, Figure 6.5).
- Active bystanders do not need to have a majority in order to provide significant protection to the out-group, they must only be sufficiently higher than the number of complicit bystanders (Section 6.3, Figure 6.6).
- Higher percentages of perpetrators in the environment yield higher death rates, with those rates dropping exponentially for $\beta = 0.2$ (Section 6.5, Figure 6.10).
- The Eris model can correctly simulate a crisis in the environment that increases persecution of the out-group (Section 6.6), or change in system-level factors of restraint in the form of a positive or negative step function (Section 6.7).

[1]Note that 15,000 days represents approximately 41 years. While genocide normally occurs over a much shorter time frame, the Eris model is designed to simulate long-term conditions in societies. Of interest here is how factors of restraint that may exist for decades influence the outcome should a crisis arise.

The following sections describe the experiments and their results in full detail.

6.2 Effects of Active Bystander Contagion and In-Group Fear

In the Eris model, there are two global conditions that can be turned on and off: Active Bystander Contagion (Algorithm 3), and In-Group Fear in the presence of local acts of violence as recorded in the environment (Algorithm 4). This experiment studies the effect of all combinations of these conditions on out-group survival. Section 4.3.1 explains the basis for including these two types of global conditions in the model. Initial runs of this sub-experiment were conducted by von Briesen et al. [51], and have been replicated here using the methodology outlined in Section 6.1. The major differences between the results reported in [51] and what is detailed below are that the Eris model allows agent Susceptibility to affect the degree of change to OrientationDifference, and results are now averaged over 100 replicates.

6.2.1 Contagion and Fear Experiment Description

In order to study the effect of all combinations of Active Bystander Contagion and In-Group Fear effects on out-group survival, we have conducted model runs for four combinations of global conditions as shown below in Table 6.1. For each condition set, the independent variable is the number of perpetrators in the environment, which ranges from 1 to 100 in steps of 1. Each of these steps is replicated 100 times, and the dependent variable measured for each is the day on which no out-group members remained in the system. In this experiment, we have examined two conditions with respect to agent Susceptibility: first, we look at the results when Susceptibility does not affect the degree of change in OrientationDifference, and then compare those results to when Susceptibility has a moderating effect (see Algorithms 2, 3, and 4).

Experimental conditions and settings:

- **Conditions:** 1-4 from Table 6.1, Agent Susceptibility moderates change to OrientationDifference (T/F)
- **Independent variable:** Number of perpetrators ([1, 100] in steps of 1)

Table 6.1: Global experimental condition sets for exploration of the effects of contagion and fear on out-group violence.

Contagion and Fear Set Number	**Active Bystander Contagion (AB)**	**In-Group Fear (IF)**
1	TRUE	TRUE
2	TRUE	FALSE
3	FALSE	TRUE
4	FALSE	FALSE

- **Dependent variable:** First day on which $OutGroupCount = 0$ or $OutGroupCount < 0.5$ (to address long-tail simulations).
- **Other experimental settings:**
  - Number of agents: 500
  - In-Group initial percentage: 70%
  - β: no function

6.2.2 Contagion and Fear Experiment Results

Figure 6.1 shows the first set of results for this experiment. For these, the dependent variable is the day on which the mean out-group count over all replicates is equal to zero. The chart legend refers to the Table 6.1 combinations by number. Note that the Susceptibility does not moderate change to Orientation difference in the left chart, and *does* have this affect in the right chart.

One can see a high level of volatility in these results, especially when compared to what von Briesen et al. found when averaging over only three replicates [51]. This is due to simulations in which one or two out-group agents remain, and conditions cause them to survive much longer than the average. These are long-tail simulations, and clearly introduce significant noise into the final result. To address this issue, we have filtered these data, taking the dependent variable to be the day on which the mean out-group count over all replicates is less than 0.5. This threshold means that about half of all 100 replicates have an out-group count of zero on that day. Figure 6.2 below shows the effect of this filtering process.

Removing the effect of long-tail simulations yields a much less volatile result. It is im-

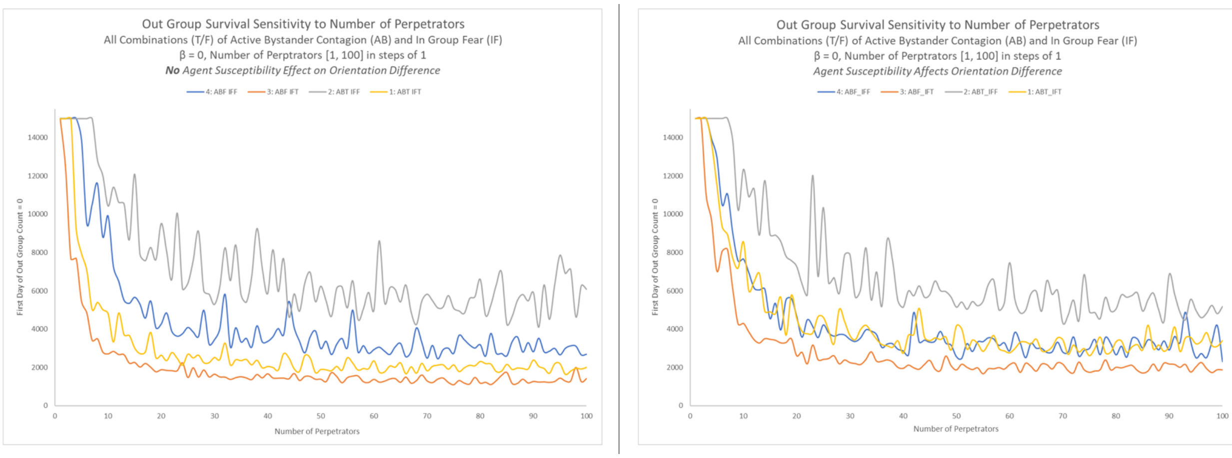


Figure 6.1: Out-Group Survival as a function of the number of perpetrators in the environment for all combinations in Table 6.1. Agent Susceptibility attribute effect on Orientation-Difference and thus BystanderChoice: Left–no effect, Right–with effect.

portant to note the degree of impact these long-tails can have, and to recall that they are typical of ABM results given the built-in stochasticity of the models. The next section provides further analysis of these results.

6.2.3 Contagion and Fear Experiment Analysis Summary

This analysis references Figure 6.2, as similarities, differences, and patterns are easier to identify in the filtered results. First, all combinations yield expected results:

- Combination 2 (AB=True, IF=False) is the best condition for out-group survival. The contagion effect causing active bystander behavior to spread to other agents provides an added protection to the out-group, reducing the death rate.
- Combination 3 (AB=False, IF=True) is the worst condition for out-group survival. Here, in-group fear causes and increased probability of an agent becoming a complicit bystander, increasing death rates for the out-group.
- Combination 4 (AB=False, IF=False) falls in between the results of Combinations 2 and 3. Here there are no external factors that influence an in-group agent's choice of BystanderType.
- Combination 5 (AB=True, IF=True) also falls in between the results of Combinations 2 and 3, and is a slightly worse outcome for the out-group when agent Susceptibility

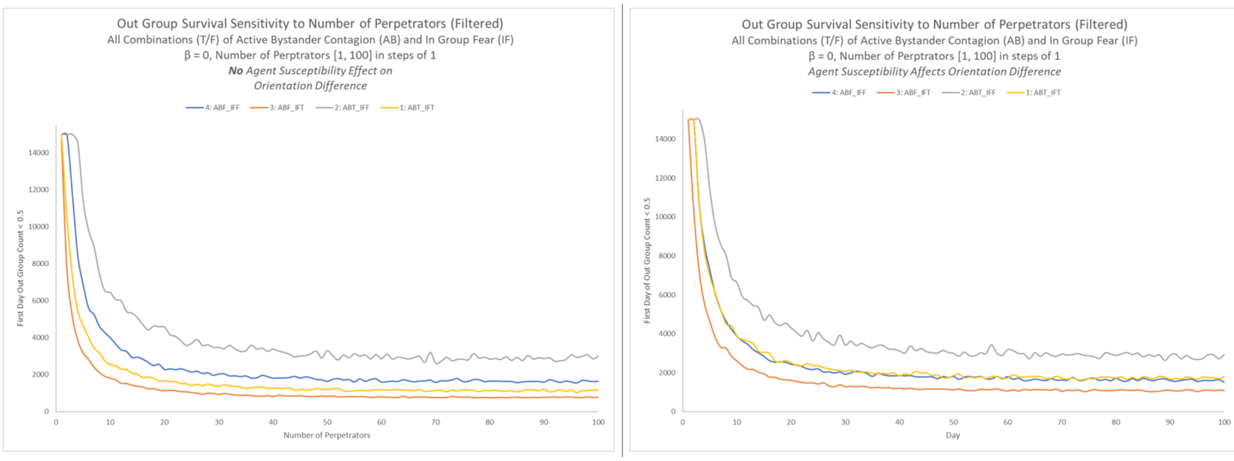


Figure 6.2: Out-Group Survival as a function of the number of perpetrators in the environment for all combinations in Table 6.1. Agent Susceptibility attribute effect on OrientationDifference and thus BystanderChoice: Left–no effect, Right–with effect. Last day of out-group survival as $OutGroupCount < 0.5$.

does not moderate the effect of these external factors on the OrientationDifference value (Figure 6.2, left). When Susceptibility *does* have a moderating effect (Figure 6.2, right), it causes Combinations 4 and 5 to yield nearly identical results.

The above results show that the Eris model is behaving as expected under the different combinations of conditions explored here. The impact of allowing Susceptibility to moderate the effect of external factors on OrientationDifference was unexpected. This causes the impact of any benefit of having an ActiveBystanderContagion to be cancelled out by also having an InGroupFear effect in Combination 1. Without the moderation of Susceptibility, there is enough of a range in Orientation difference to show a slightly more sustainable environment for the out-group with Combination 4.

6.3 Sensitivity to Factors of Restraint (β)

In their recent work, von Briesen et al. explored the effect of global factors of restraint (β) on violence against the out-group [51]. Again, the major differences between the results reported in [51] and what is detailed below are that the Eris model allows agent Susceptibility to affect the degree of change to OrientationDifference, and results are averaged over 100 replicates. As with the experiment described in Section 6.2.1, we examine two conditions with respect to agent Susceptibility: with and without the moderating effect of Susceptibility on OrientationDifference (see Algorithm 2).

6.3.1 β Sensitivity Experiment Description

The purpose of this experiment is to determine how sensitive the Eris model is to the β function simulating factors of restraint. Again, all but the final result compare the sensitivity to β with and without agent Susceptibility moderating the global function's effect on OrientationDifference. We have performed these experiments under Condition 4 from Table 6.1 (AB=False, IF=False). This is an appropriate starting point for β sensitivity analysis, as it isolates the effect of β on the outcome.

Experimental conditions and settings:

- **Conditions:** Condition 4 (AB=False, IF=False)
- **Independent variable:** Day
- **Dependent variable:** Group Counts
- **Other experimental settings:**
  - Number of agents: 500
  - In-Group initial percentage: 70%
  - Number of perpetrators: 50
  - β:
    - $[0.09, 0.49]$ (without agent Susceptibility effect on OrientationDifference)
    - $[0.09, 0.85]$ (with Susceptibility effect)
    - In steps of 0.02

6.3.2 β Sensitivity Experiment Results

Figure 6.3 compares the sensitivity of out-group survival to β, with the left chart showing results when there is no effect of agent Susceptibility on OrientationDifference, and the right including that effect. When there is no effect, the out-group receives strong protection due to the presence of active bystanders as β approaches 0.3. In contrast, when accounting for agent-level, individual variations in Susceptibility to the influence of system-level conditions, β nearing 0.3 is not sufficient to prevent violence. While out-group survival improves with increasing β, there is still a significant decline over time.

Figure 6.4 shows the results for a wider range of β when allowing agent Susceptibility to affect OrientationDifference. Here, the out-group does not experience significant protection

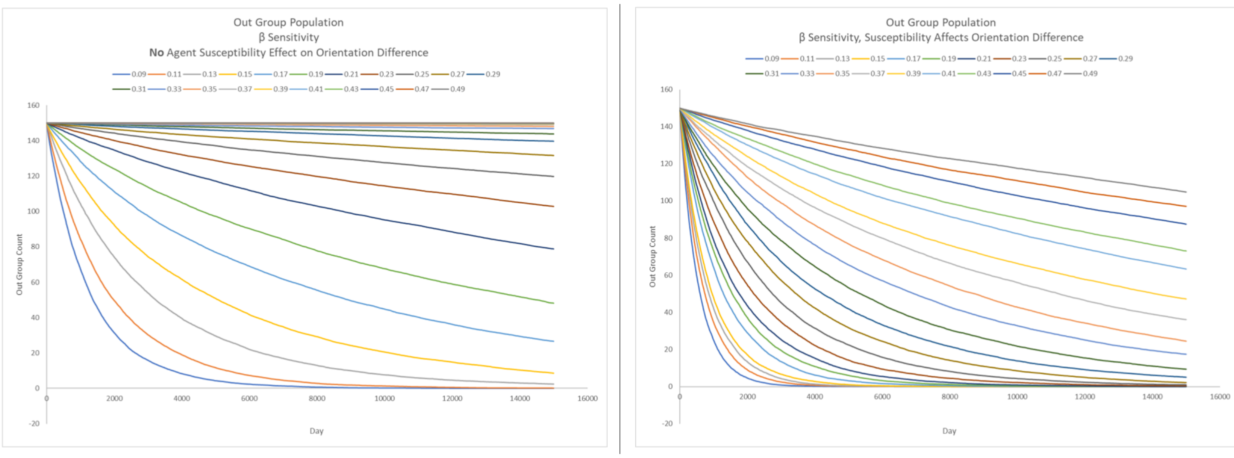


Figure 6.3: Out-group population sensitivity to factors of restraint in the society (β). Agent Susceptibility attribute effect on OrientationDifference and thus BystanderChoice: Left–no effect, Right–with effect.

due to the presence of β until it exceeds 0.6. This result makes sense, as individual agent Susceptibility is initialized according to a normal distribution (see Table 4.2), which reduces the effect of β on every individual agent according to how susceptible they are to change. This is quite different than allowing the full value of β to impact the OrientationDifference value in Algorithm 2.

Figure 6.5 provides a clearer quantification of the effect accounting for agent Susceptibility has on out-group death rate for different values of β by visualizing the slopes of linear lines in the β sensitivity plots.[2] Of note here is that in both cases, higher β has an exponentially beneficial effect on out-group protection from violence. However, it is clear that when accounting for the effect of agent Susceptibility to change, β must be greater than 0.8 for full out-group protection under the conditions detailed in Section 6.3.1. This means that any individual agent must have an 80% chance of becoming an active bystander due to the type of society in which they live. Recall that in addition to 500 civilian agents, there are 50 perpetrators in this environment, which is a high percentage and has its own impact on the results. Section 6.4 explores how these dynamics change with respect to population size for

[2]Of note in Figure 6.5 is the greater smoothness of the left curve in which Susceptibility does not affect OrientationDifference. Allowing agent Susceptibility to influence the outcome increases the level of stochasticity in the model, causing increased variation in the curve to the right. As such, future work with this model setting should require more than 100 replicates in order to determine a stable average value for results.

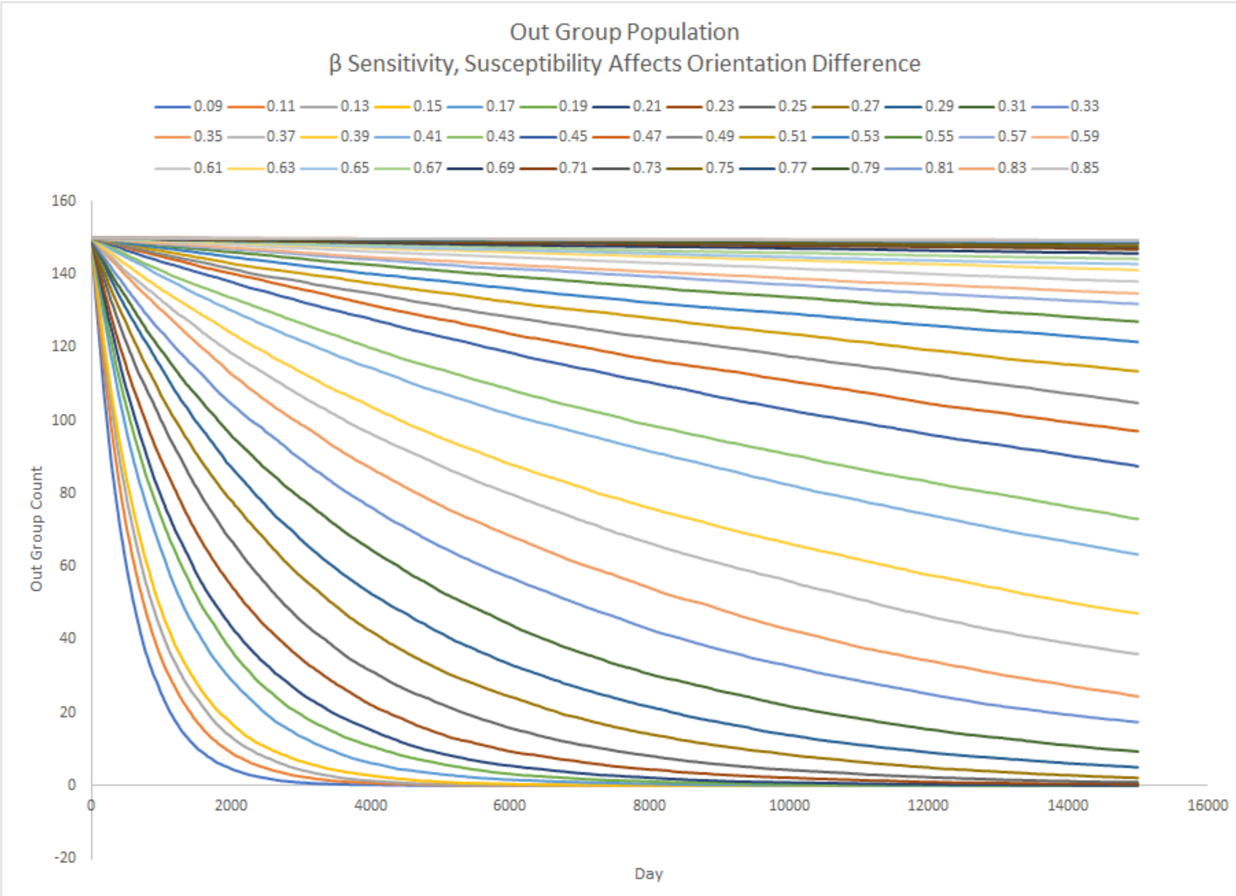


Figure 6.4: Out-group population sensitivity to factors of restraint in the society (β). Agent Susceptibility attribute affects Orientations Difference. β has expanded range: $[0.09, 0.85]$.

different levels of perpetrators in the environment.

Figure 6.6 presents one final result for this exploration into β sensitivity. Behind the scenes of out-group persecution by perpetrators are bystanders who influence the outcome. As previously discussed, higher levels of active bystanders yields higher levels of out-group protection, and higher values of β means the society itself has greater restraints against persecution of minority groups. The left chart in the figure is for $\beta = 0.09$, and the right is for $\beta = 0.29$. Note that for the higher β the number of active bystanders is nearly twice that in the alternate scenario. In fact, for $\beta = 0.09$ the out-group is nearly annihilated by day 5000, while there is a very low death rate for the higher β value. Of particular interest is that for $\beta = 0.29$, the number of passive bystanders is nearly twice that of the active type. This

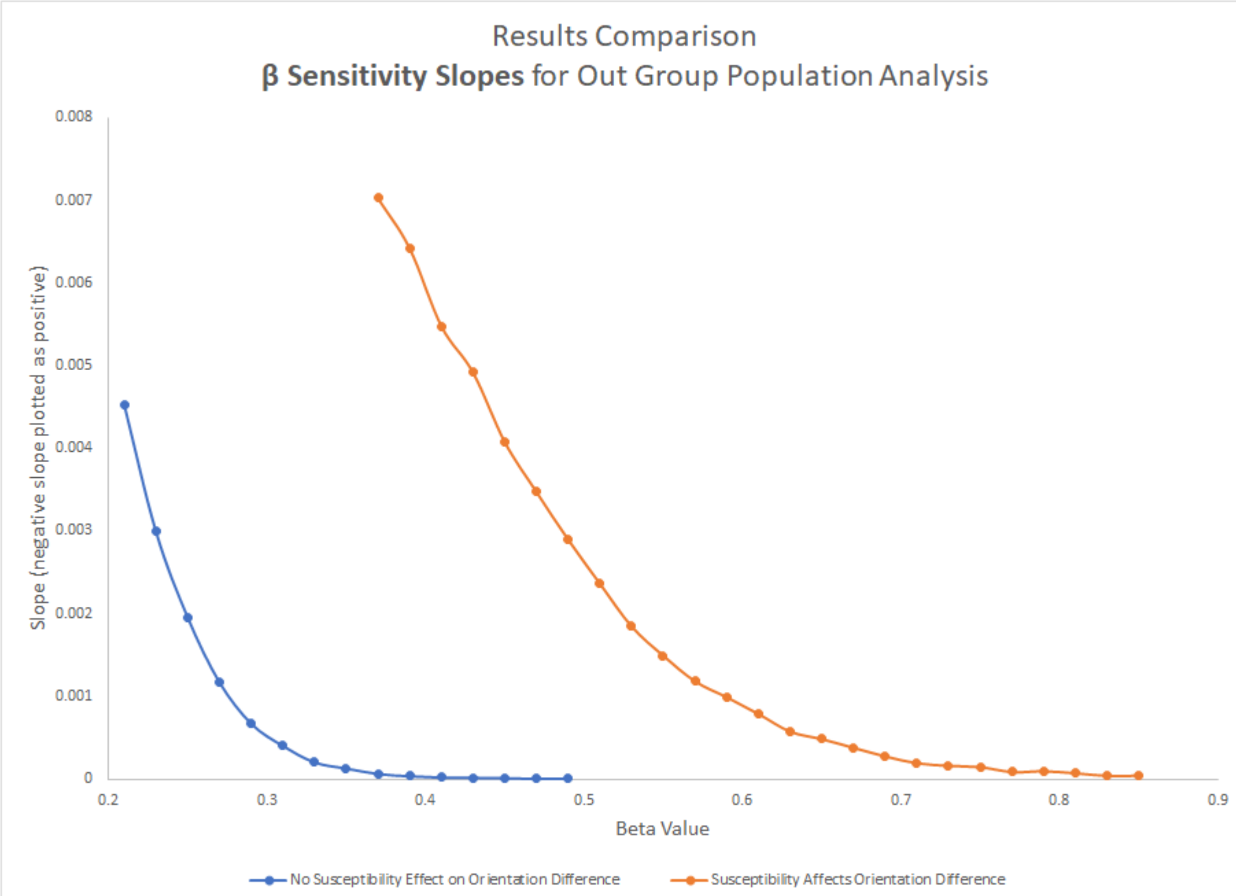


Figure 6.5: Comparison of slopes of linear β sensitivity plots with and without agent Susceptibility effect on OrientationDifference (from Figures 6.3 and 6.4. Note that negative slope is plotted as positive to better visualize decreasing death rate.

means that active bystanders do not need to be a majority in order to protect the out-group, they only need to be sufficiently higher than the number of complicit bystanders present.

6.3.3 β Sensitivity Experiment Analysis Summary

The above results lead to three major findings with respect to out-group population sensitivity to β:

- Figure 6.5: Accounting for individual, agent-level differences in Susceptibility to change has a dramatic impact on how well factors of restraint in the form of the β function protect the out-group from persecution. For the experimental conditions outlined in Section 6.3.1, more than 80% of the in-group must have a high probability of being

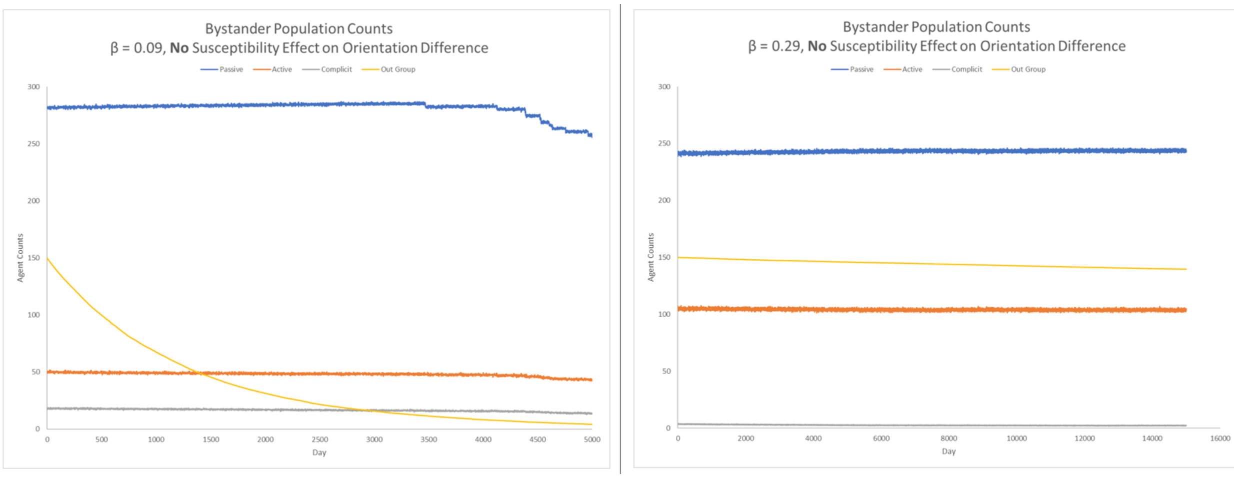


Figure 6.6: Sample bystander counts compared to out-group population for two different values of β. Here there is no effect of agent Susceptibility on OrientationDifference calculation in Algorithm 2.

an active bystander in order to protect the out-group given the high percentage of perpetrators present in the environment.

- Figure 6.5: β has an exponentially beneficial effect on out-group death rates. This is the same finding in [33]. As such, policy initiatives that can move a society toward inclusion can have exponentially beneficial effects for out-group members under certain circumstances.
- Figure 6.6: Active bystanders do not need to be a majority in the environment in order to provide sufficient protection to the out-group. With sufficiently high β values, passive bystanders can remain a majority and the number of active bystanders will be sufficient to prevent high death rates.

These findings show the importance of system-level factors of restraint (β) for protecting an out-group population. The next experiment explores how population size and percentage of perpetrators present influence these results.

6.4 Sensitivity to Population Size

In order to better understand and verify the results the Eris model can provide, we next explore if and how the model dynamics change with respect to population size. Prior experiments in this chapter kept the population size fixed at 500, with the in-group representing 70% of civilians. Here, population size ranges from 500 to 5000 in steps of 500. As for the number of perpetrators (numPerps), we explore two scenarios: numPerps = 50, and numPerps = 10% of the total population size.

6.4.1 Population Size Experiment Description

This is a very high-level approach to understanding the influence population size has on model dynamics. First, this experiment takes Condition 4 as a baseline (AB=False, IF=False), and has no factors of restraint influencing the outcome ($\beta = 0$). The purpose of choosing these settings is to determine the effect of changing population size in the absence of any other factors. Note that in this case, agent susceptibility does not affect the OrientationDifference calculation; however, that ultimately has no bearing given there are no factors of escalation or restraint present.

Experimental conditions and settings:

- **Conditions:** Condition 4 (AB=False, IF=False), Individual agent susceptibility does *not* affect OrientationDifference calculation.
- **Independent variable:** Out-Group Count
- **Dependent variable:** Day
- **Other experimental settings:**
  - Number of agents: [500, 5000] in steps of 500
  - In-Group initial percentage: 70%
  - Number of perpetrators:
    - Set 1: numPerps = 50
    - Set 2: numPerps = 10% of the total population size
  - β: no function

6.4.2 Population Size Experiment Results

Figure 6.7 below shows the results of the experiment described above. First, the general dynamics of out-group death do not change given increased civilian population size. This is a good indication that the model is stable for different numbers of agents. More surprisingly, the difference between 50 perpetrators as a fixed value, and the perpetrator count being 10% of the civilian agent count appears to be relatively insignificant. This is unexpected, as it seems that a higher percentage of perpetrators in the environment should have a stronger effect.

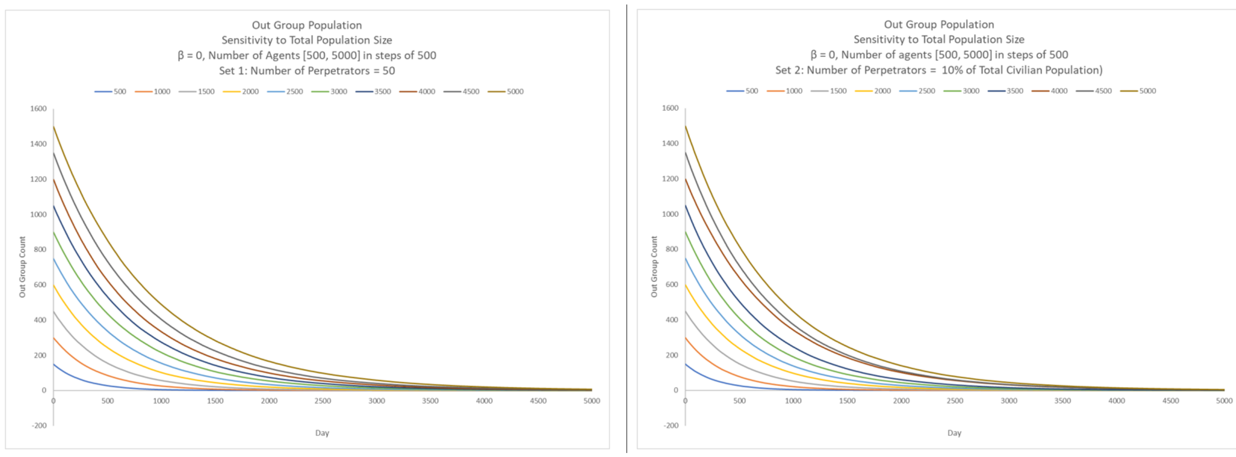


Figure 6.7: Sensitivity of out-group population to the overall population size under two conditions: number of perpetrators is fixed at 50, and number of perpetrators is 10% of the civilian population count.

In order to examine this outcome more carefully, Figure 6.8 compares the results under both conditions for 5000 agents. Again, while the out-group does experience a lower death rate, the difference when perpetrators are around 1% versus 10% of the total number of civilians is not dramatic.

In order to understand this result, we return to Algorithm 5, which describes the rules for deciding if an out-group agent will survive when in the presence of a perpetrator. The probability of death for this agent is lower for high ratios of active bystander neighbors, and higher for high ratios of complicit bystander neighbors. The density of perpetrator agents does not affect this probability, reducing the effect their relative numbers have on the death rate. This is clearly unrealistic, and future versions of the Eris model should adjust the logic of Algorithm 5 to better capture expected outcomes.

6.4.3 Population Size Experiment Analysis Summary

This experiment yielded two major findings:

- Overall dynamics for out-group survival are similar for higher and lower civilian populations. The death rate is roughly exponential when there are no factors of restraint

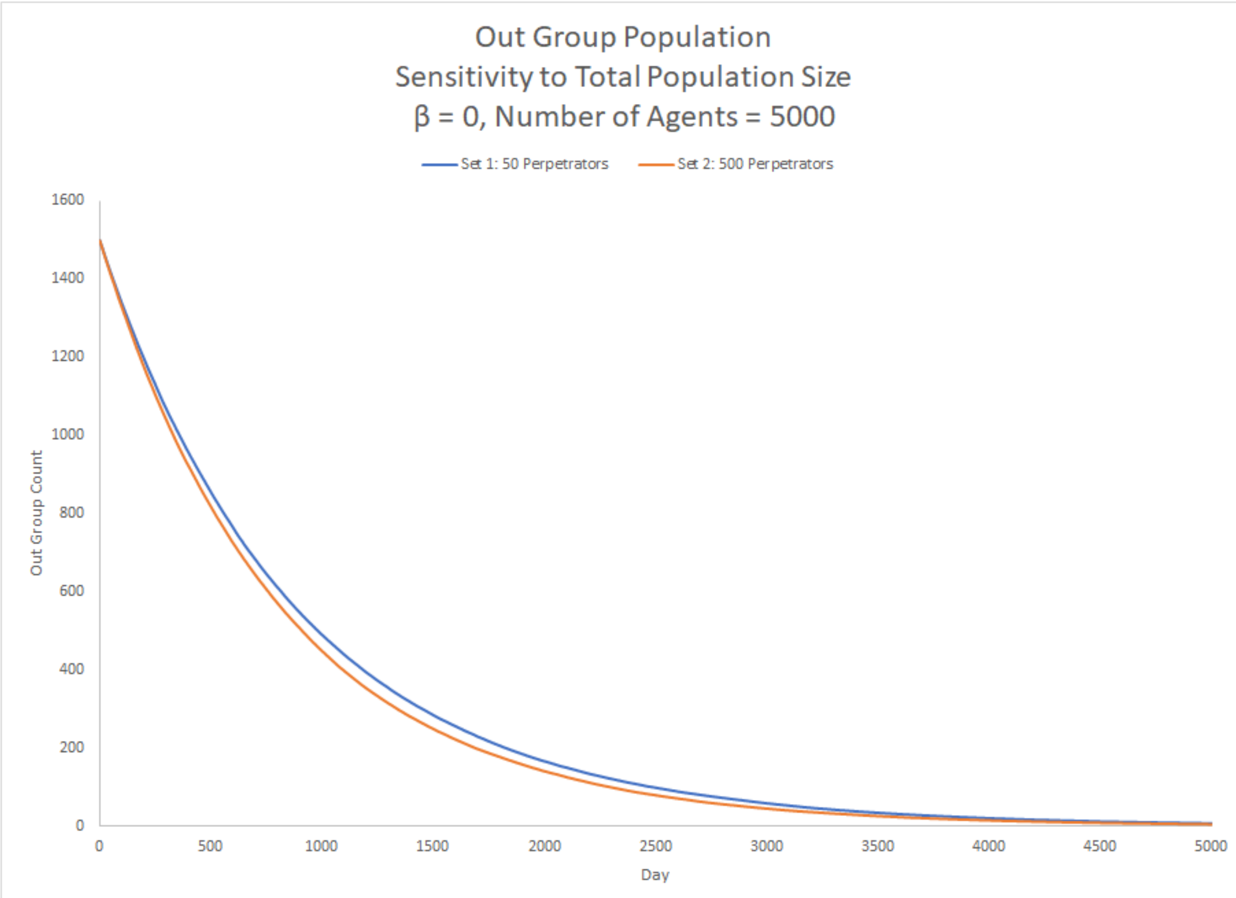


Figure 6.8: Sensitivity of out-group population to the overall population size for 500 agents under two conditions: number of perpetrators = 50, and number of perpetrators = 10% of the civilian population count.

($\beta = 0$).

- Higher perpetrator density does not significantly impact death rates due to logic of Algorithm 5. Future versions of the Eris model should adjust this logic to allow a higher probability of death given higher perpetrator density.

6.5 Out-Group Sensitivity to Perpetrator Count

This experiment follows from the previous explorations of out-group population sensitivity to β (Section 6.3), and to population size (Section 6.4). Note that in this experiment, agent Susceptibility continues to moderate changes to OrientationDifference.

6.5.1 Out-Group Sensitivity to Perpetrator Count Experiment Description

The motivation for this experiment is to determine the effect the number of perpetrators has on the outcome for different, constant β values across two different population sizes. The results of the β sensitivity experiment presented in Section 6.3 showed that an environment with 500 agents and 50 perpetrators required a very high β value to restrain violence. It is unlikely that 10% of the population would be perpetrators, therefore this experiment explores a range of perpetrator density from 0.2% - 10% of the total civilian population count. For two values of β, we also look at the effect of population size on the results for 500 and 1000 agents.

Experimental conditions and settings:

- **Conditions:** Condition 4 (AB=False, IF=False), agent Susceptibility affects OrientationDifference in all the following experiments:
  - Set 1: $\beta = 0.5$, Number of agents = 500, Number of perpetrators [5, 50] in steps of 5
  - Set 2: $\beta = 0.5$, Number of agents = 1000, Number of perpetrators [10, 100] in steps of 10
  - Set 3: $\beta = 0.2$, Number of agents = 500, Number of perpetrators [5, 50] in steps of 5
  - Set 4: $\beta = 0.2$, Number of agents = 1000, Number of perpetrators [10, 100] in steps of 10
- **Independent variable:** Day
- **Dependent variable:** Out-Group Count
- **Other experimental settings:**
  - Number of agents: See Conditions section above for ranges by experiment.
  - In-Group initial percentage: 70%
  - Number of perpetrators: 10% of civilian agent population. See Conditions section above for ranges by experiment.
  - β: Sets 1 and 2: $\beta = 0.5$, Sets 3 and 4: $\beta = 0.2$

6.5.2 Out-Group Sensitivity to Perpetrator Count Experiment Results

Figure 6.9 below shows the results for this experiment with $\beta = 0.5$. The left chart shows that in a society that has a 50% higher probability of an agent becoming an active bystander, the out-group population does decline in a linear fashion, with a steeper decline for higher numbers of perpetrators. Notably, for the same simulation with 1000 agents (right), even

100 perpetrators can achieve very little persecution of the out-group, with approximately 2-3 deaths over 15,000 days. This result is quite unexpected, and appears to be caused by how Algorithm 5 calculates the probability of out-group death. Again, additional perpetrators in the vicinity of an out-group agent do not increase the probability of death. However, the increase in density means that there will be an increase in active bystanders, particularly with a higher value of β. As stated in Section 6.4.3, future versions of the model should account for increased risk to the out-group given a higher density of perpetrators.

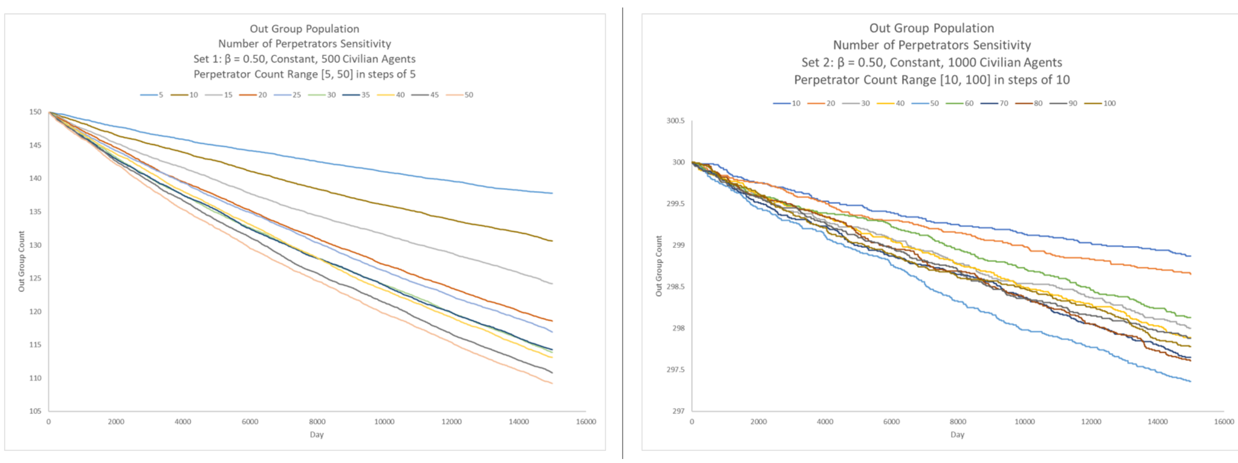


Figure 6.9: Out-group sensitivity to number of perpetrators with $\beta = 0.5$, for 500 (left) and 1000 (right) agents.

Figure 6.10 shows the results of this experiment for β = 0.2. These results confirm that for this low value of β, the out-group will experience extreme levels of violence and persecution as the number of perpetrators increases. As with the previous result, the effect is muted for the larger overall population size; however, the difference is not as dramatic as when β = 0.5 because there are overall fewer active bystanders given the lower β value.

6.5.3 Out-Group Sensitivity to Perpetrator Count Experiment Analysis Summary

We have two findings for this experiment:

- Higher percentages of perpetrators in the environment yield higher death rates, with those rates dropping exponentially for β = 0.2
- As was the case in the previous experiment, higher population density causes lower

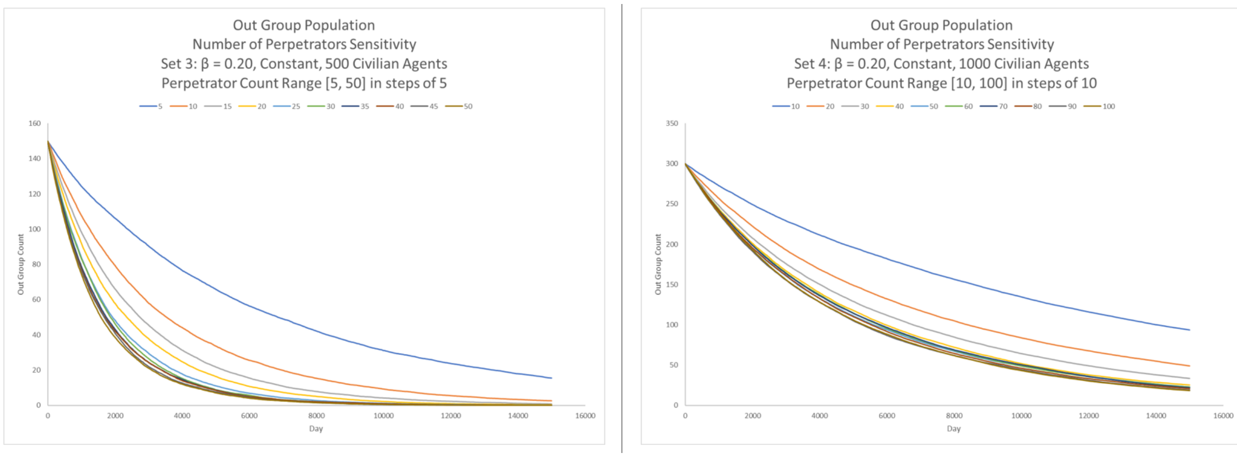


Figure 6.10: Out-group sensitivity to number of perpetrators with $\beta = 0.2$, for 500 (left) and 1000 (right) agents.

death rates. Future versions of the Eris model should adjust the logic of Algorithm 5 to allow a higher probability of death given higher perpetrator density.

The next two experiments employ Eris model functionality implemented to allow for empirical validation. To achieve this type of validation, the model must be calibrated to simulate a specific scenario. In order to maintain Eris' generalized framework, this is achieved in two ways as explained in Section 4.2.2: through modeling the emergence of a crisis by increasing the number of perpetrators in the environment, and through modeling β as a positive or negative step function to simulate a change in the tone of elite communication to civilians. Note that for all remaining experiments, agent Susceptibility moderates the effect of conditions on Orientation Difference (see Algorithms 1, 2, 3, and 4).

6.6 Triggering a Crisis

This experiment explores the dynamics of a crisis occurring in the Eris environment. This experiment is conducted with four sets of conditions detailed below. In order to explore the impact ActiveBystanderContagion has on the overall outcome in the presence of a crisis, two sets allow for this contagion effect and two do not. Again we examine the sensitivity of out-group population to a range of β values.

6.6.1 Crisis Trigger Experiment Description

Experimental conditions and settings:

- **Conditions:** Agent Susceptibility affects OrientationDifference in all the following experiments:
  - Set 1: Condition 4 (AB=False, IF=False), Perpetrators increase from 0 to 50 on day 5000, β ranges from [0, 0.95] in steps of 0.05
  - Set 2: Condition 2 (AB=True, IF=False), Perpetrators increase from 0 to 50 on day 5000, β ranges from [0, 0.95] in steps of 0.05
  - Set 3: Condition 4 (AB=False, IF=False), Perpetrators increase from 5 to 50 on day 5000, β ranges from [0, 1] in steps of 0.05
  - Set 4: Condition 2 (AB=True, IF=False), Perpetrators increase from 5 to 50 on day 5000, β ranges from [0, 1] in steps of 0.05
- **Independent variable:** Day
- **Dependent variable:** Out-Group Count
- **Other experimental settings:**
  - Number of agents: 500
  - In-Group initial percentage: 70%
  - Number of perpetrators: See Conditions section above for ranges by experiment
  - β: See Conditions section above for ranges by experiment

6.6.2 Crisis Trigger Experiment Results

Figure 6.11 below shows the results with and without ActiveBystanderContagion for perpetrators increasing from 0 to 50 on day 5000. In this hypothetical scenario, the out-group experiences no threat for some time, and has a stable population. When the number of perpetrators in the environment spikes to 50, death rates are exponential for low values of β and linear for higher values, eventually reaching full out-group protection in environments with extremely strong restraining factors. This outcome is expected, as is the beneficial effect of an ActiveBystanderContagion. While this contagion effect does not completely stop violence, it clearly slows the death rate, allowing more protection at lower β values than when it is not present.

Figure 6.12 examines the dynamics of a crisis, with and without ActiveBystanderContagion, when the simulation includes perpetrators from the first day. Here we can see that death rates rise with the perpetrator count on day 5000, with an ActiveBystanderContagion providing additional protection when it is present.

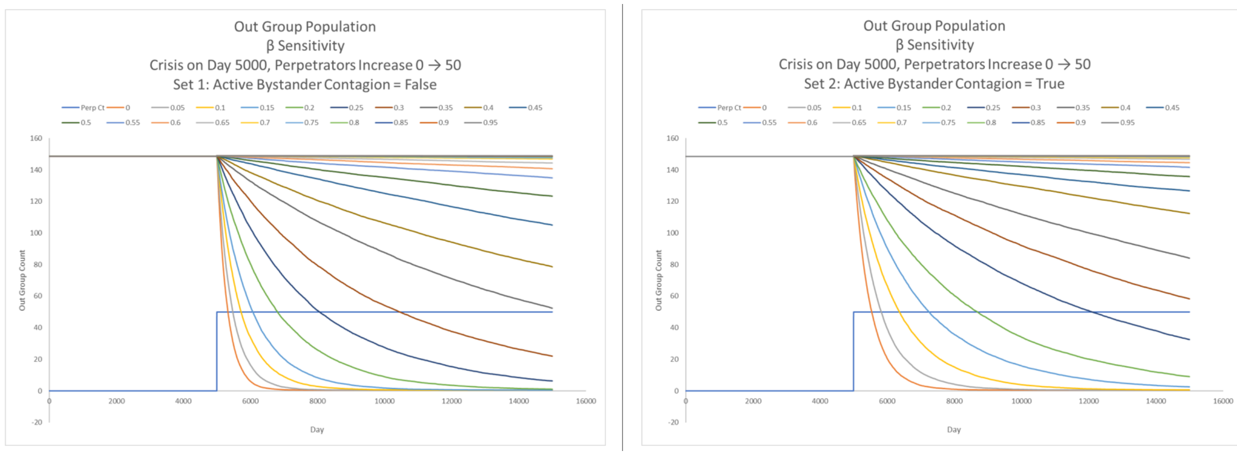


Figure 6.11: Crisis trigger experiment with initial Number of Perpetrators = 0. Determines out-group population sensitivity to β with (left) and without (right) ActiveBystanderContagion.

The above results indicate the the model is performing as expected.

6.6.3 Crisis Trigger Experiment Analysis Summary

This experiment yielded only one significant finding:

- In the event of a crisis, an ActiveBystanderContagion effect reduces the out-group death rate, meaning that lower values of β provide higher levels of protection.

While the above finding is expected, the experiment does show the usefulness of the Eris model for simulating a crisis in the environment. This is an extremely important feature, as it allows for simulation of changing environmental conditions that put the out-group at risk of persecution.

6.7 β as a Step Function

The final experiment in this chapter explores the dynamics of β as a step function. Previously, β has been constant throughout the entirety of every experiment; however, this does not realistically capture the potential for changing societal conditions. The elite power structure of a society may become more or less extreme over time, sometimes with a significant change occurring quickly. This can happen with variations in the degree of threat the elite perceives as present, or through an event like an election or coup. The change can be positive

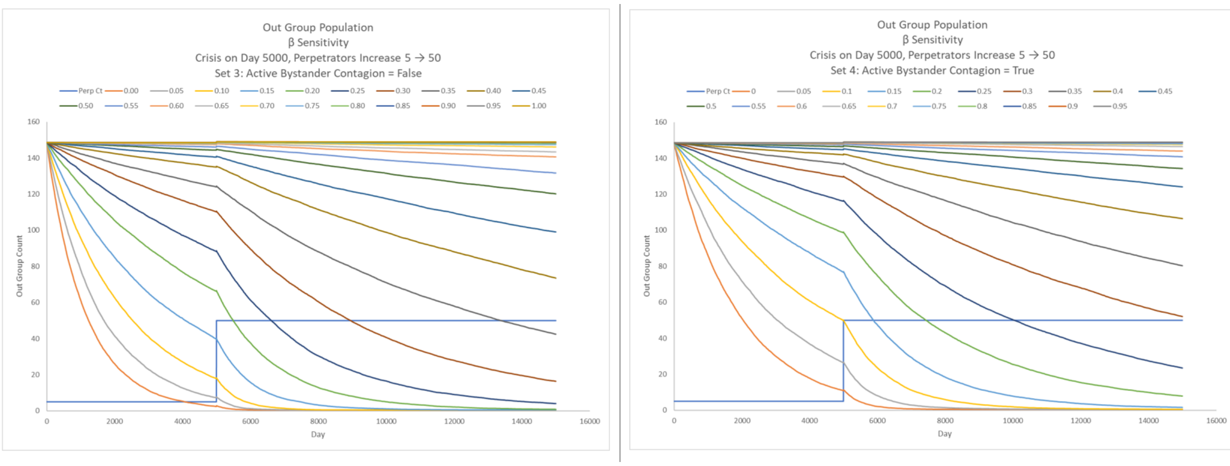


Figure 6.12: Crisis trigger experiment with initial Number of Perpetrators = 5. Determines out-group population sensitivity to β with (left) and without (right) ActiveBystanderContagion.

or negative, increasing or decreasing restraints against out-group persecution.

6.7.1 β Step Function Experiment Description

Experimental conditions and settings:

- **Conditions:** Condition 4 (AB=False, IF=False), agent Susceptibility affects OrientationDifference in all the following experiments:
  - Set 1: Positive step on day 500. Initial β = 0, β step [0.05, 0.95] in steps of 0.05
  - Set 2: Negative step on day 500. Initial β = 1, β step [0.05, 0.95] in steps of 0.05
- **Independent variable:** Day
- **Dependent variable:** Out-Group Count
- **Other experimental settings:**
  - Number of agents: 500
  - In-Group initial percentage: 70%
  - Number of perpetrators: 50
  - β: Step function, see Conditions section above for details

6.7.2 β Step Function Experiment Results

Figure 6.13 shows the results of Set 1 of this experiment, in which the β function increases on day 500, introducing rising factors of restraint into the society. Here, the initial value of β is 0, resulting in a very high death rate until β increases on day 500. As expected, very high values of β offer protection for the remaining out-group members. The chart on the left shows the full results, while that on the right zooms in to the area closer to day 500.

It is difficult to determine exactly how the death rate changes with respect to β in View 2

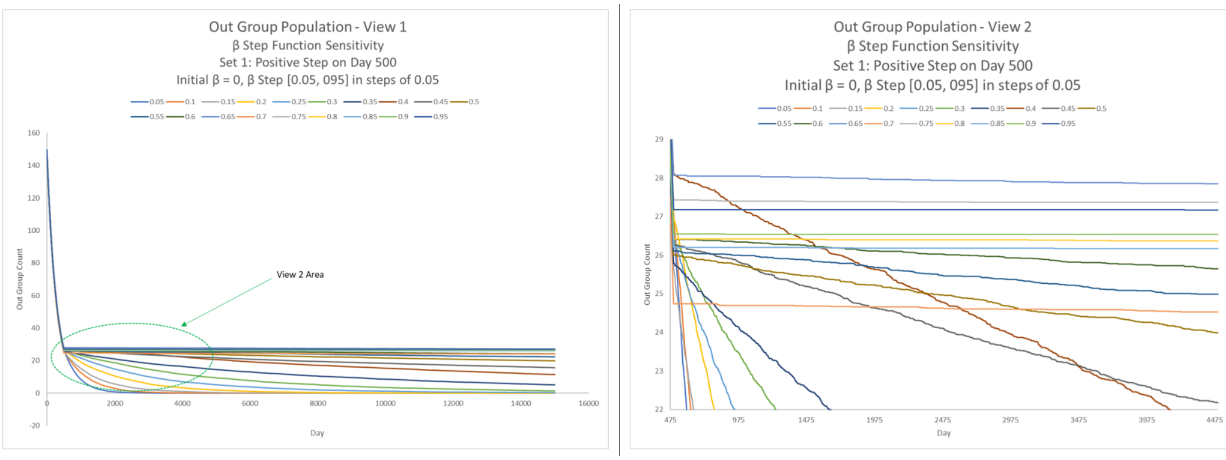


Figure 6.13: Out-group population sensitivity for positive β function with initial value of 0.

of Figure 6.13. Figure 6.14 below displays these results through a different lens: the slope of each line as it corresponds to β. We can see here that the level of protection offered to the out-group increases exponentially with increasing β, and that what appears to be a rather noisy result in Figure 6.13 is actually a smooth, exponential decline in death rate.

Figure 6.15 shows the results for a negative β step function (Set 2). In this case, the initial value of β is 1, providing full protection to the out-group. As expected, View 1 (left) reveals an increasingly worse situation for the out-group as β decreases. However, View 2 contrasts with that for the case of a positive step function, with it clearly showing a logical pattern of increase in death rates in the first days after the change. This difference in dynamic change close to the point of decrease in β is due to the fact that a negative step function introduced into a society that is completely at peace is *starting* a trend, rather than reversing one, as was the case with the positive step from Figure 6.13.

6.7.3 β Step Function Experiment Analysis Summary

There is one finding for this experiment:

- Greater degrees of change in β at the time of a positive or negative step have an exponential effect. For a positive step, high levels of increase have an exponentially beneficial effect on out-group survival. For a negative step, the opposite is true, high levels of decrease lead to exponentially increasing out-group death rates.

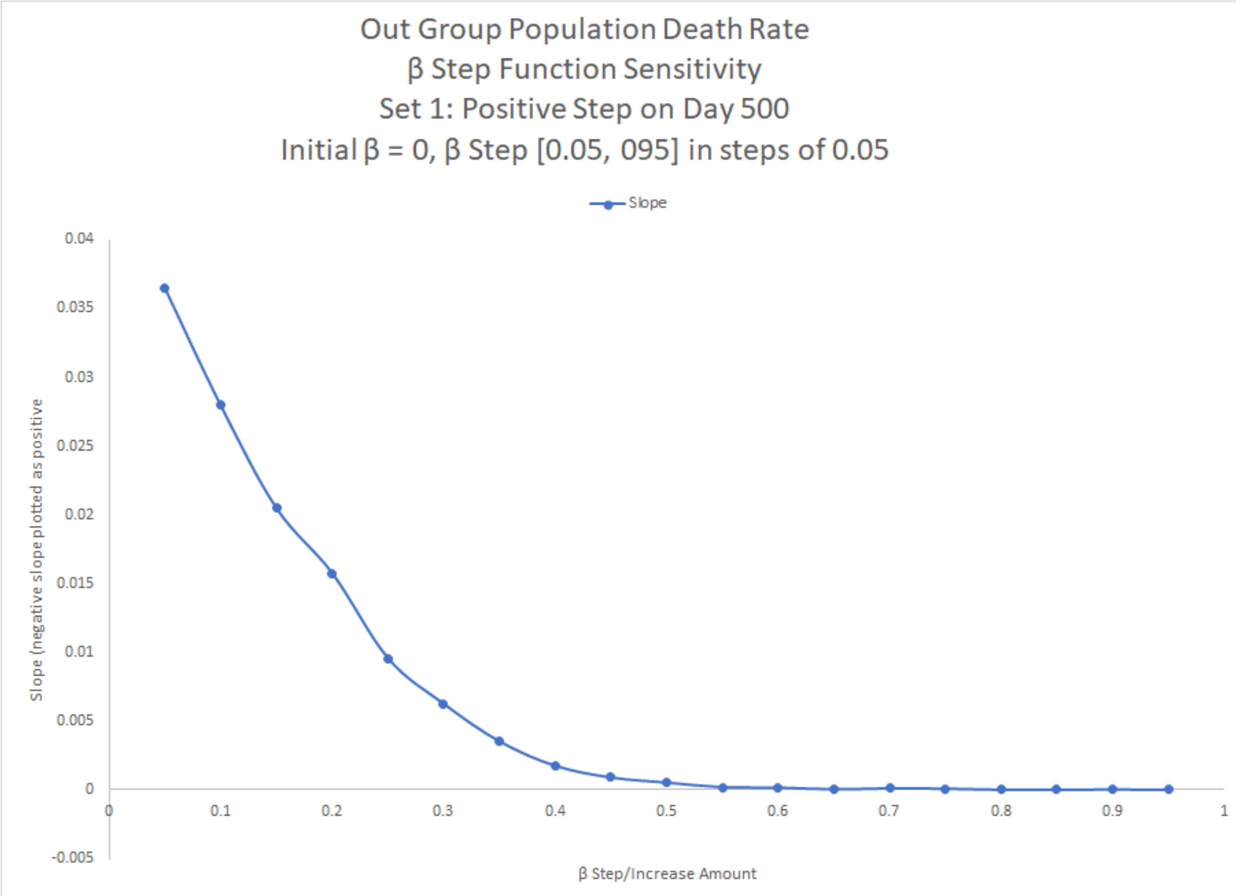


Figure 6.14: Slopes of lines shown in Figure 6.13 View 2.

6.8 Contributions and Conclusions

Section 1.2 outlined three research questions. The work of this set of experiments with the Eris model was conducted to answer Question 1 from Section 1.2.1, and part of Question 3 from Section 1.2.3. Question 1 asks what is the optimal, small and efficient model of identity-based conflict, and the journey from the literature review of history, social science, complex systems, and computational social science yields the Eris model framework as an answer. The framework is based on the simple human motivation for self- and species-preservation, making it small and efficient by reducing the complexity of agents to a simple and fundamental state. This choice of attributes also makes the model flexible across scenarios, as human beings share in these basic motivations regardless of time and place.

Question 3 asks if the model is useful and understandable to social science researchers,

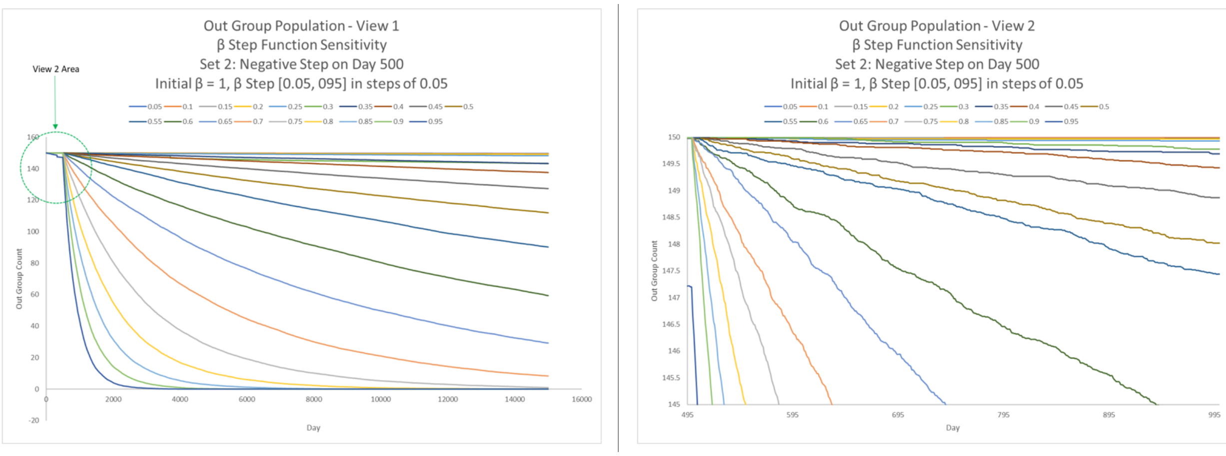


Figure 6.15: Out-group population sensitivity for negative β function with initial value of 1.

policy makers, and others. The experimental results presented in this chapter are a partial answer to this question. These show that the Eris model can successfully capture a variety of dynamics with few model settings. The most significant feature is Eris' ability to show the effect of factors of restraint in the form of a global function (β) on out-group survival when their identity is under threat.

Contributions. In sum, the Eris model and its experiments as presented in this chapter yield the following contributions:

1. The Eris model framework offers an artificial world in which to explore the difficult problem of genocide. It achieves this by taking an approach informed by complexity theory and complex adaptive systems, which results in a simple model that provides useful quantitative measures.
2. This research develops a methodology to address long-tail simulations causing increased volatility in some results that impacts interpretability.

However, parts of Research Question 3 remain unanswered. Can the Eris model be validated against a real-world scenario? Can it model the emergence or *non*-emergence of genocide similar to historic examples? Chapter 7 details the use of the Eris model to answer these questions and achieve the goal of contributing a generalized model that can also be used to reproduce real-world dynamics from different cases.

CHAPTER 7: EXPERIMENTAL SET 2 – EMPIRICAL VALIDATION THROUGH COMPARISON OF OUTCOMES ACROSS CASES

This chapter presents the first transitional steps from using the Eris model for generalized scenarios as described in Chapter 6, to empirical validation through the simulation of historic cases. Returning to Figure 3.1, the goal here aligns with the "Other Forms of Validation" portion of the spectrum, as it is between the general and specific. This approach to validation is heavily grounded in the work of Scott Straus [7, 8]. Straus' research examines "broad patterns in large-scale violence," particularly how those patterns differ between positive and negative cases [8, p. 90]. The Eris model is designed to reproduce these broad patterns in a flexible manner so the user can easily simulate different scenarios. We begin with a description of the validation experiment and its methodology (Section 7.1), followed by early results of the empirical validation (Section 7.2), analysis of those results (Section 7.3), and an overview of contributions along with concluding thoughts (Section 7.4).

7.1 Experimental Set 2–Description & Methodology

Hypothesis. If the Eris model is calibrated to reproduce the death rate of the 1994 Rwandan genocide, and then the β value associated with that outcome is increased by the magnitude of difference in societal restraints found for Côte d'Ivoire shown in Table 5.5, we hypothesize that this increased β will greatly stem or completely prevent violence against the out-group in the Rwandan scenario. If this hypothesis is found to be true, increased ecosystem levels in elite speech is a potential explanation for why genocide may or may not occur in scenarios where it is expected.

Chapter 4 outlined the components and logic behind the Eris model, and was followed by Chapter 5, a description and analysis of data from Rwanda and Côte d'Ivoire intended for use in the empirical validation process. This section explains how the model and data are linked, using the findings presented in Chapter 6 to inform the process.

7.1.1 Linking the Model and Data

Section 4.6.1 introduced the β function as the Eris model's abstraction of system-level factors of restraint against out-group persecution. This function is designed to be a link between data that quantifies these restraints for a given society and the model environment (see Figure 4.1). Chapter 5 explains the use of sentiment analysis on presidential speech data from Rwanda and Côte d'Ivoire to quantify the ecosystem motivational orientation of the elite in these two societies prior to their respective crises. Table 5.5 summarizes the comparison of mean ecosystem values for both cases, and we will use the differences between these mean scores to inform the model's β function as follows:

1. The model will be calibrated to reproduce the the Rwandan genocide.
2. The β value associated with reproduction of the death rate in Rwanda will be taken as a baseline.
3. Using the results from Table 5.5, the β function will be increased to align with the findings from Côte d'Ivoire to determine if this level of restraining factors is enough to prevent a genocide in the Rwandan scenario. If this increased β value prevents violence, this then provides a potential explanation for how the Ivorian crisis of the early 2000s did not devolve into genocide.

7.1.2 Additional Scenario Specific Data for Case 1: Rwanda

In addition to the use of data to inform the β function, we require specific information about the Rwandan population in 1994 for model calibration. Note that the values listed are approximate:

- Total population: 7,000,000 [8, p.274]
- Hutu (in-group) percentage: 85% (5,950,000) [28]

- Tutsi (out-group) percentage: 14% (980,000)[1]
- Number of perpetrators: between 175,000 and 210,000 (midpoint is 192,500, or 3.2% of Hutu population) [140]
- Death rate: between 800,000 and 1,000,000 in 100 days (82% of all Tutsis for lower estimate, or 8000 per day) [14, p. 360]

The above data informs model parameters and validation as follows:

1. Model will be initialized with a synthetic population in which 85% represent the in-group. Note that we do not account for the Twa population and allow 15% of the total to represent the Tutsi out-group at this stage.
2. Perpetrator count will be set at 3.2% of the in-group total. Note that this is not an exact representation of the true perpetrator/in-group ratio, as there were many civilian participants in the genocide. The value serves as a rough estimate at this stage.
3. Model will run until 82% of the out-group has died and been removed from the environment.

7.1.3 Revised Model: Eris2

Chapter 6 presented the results of model exploration for a general scenario, revealing the capabilities and limitations of Eris for studying identity-based conflict. Before proceeding, it is important to highlight two particular items from these results which will inform model adjustments required to achieve empirical validation:

- Simulations in Chapter 6 were run for 15,000 days, or approximately 41 years. While genocides do not occur over that long a time frame, Straus' work suggests that what happens in societies long before a crisis is important. As such, the first version of the Eris model is designed to be stable over many decades.
- The results presented in Sections 6.4 and 6.5 underscore the limitation of the model framework when simulating higher perpetrator density in the environment. Higher density should lead to higher risk to the out-group, and addressing this was left for

[1]Note that the Twa represent 1-2% of the population [14, p. 350][28]

future model versions.

Initial test runs showed that the Eris model could not be calibrated to reproduce the Rwandan death rate even under the most extreme conditions of $\beta = 0$ and "InGroupFear" acting as a micro-level factor of escalation. Taking into consideration the key points presented above, we found that the model required the following adjustments to the determination of the probability of out-group agent death:

1. Experiments are run with Condition 3 (AB=False, IF=True) from Table 6.1. This allows the simulation to better capture the intensity of the Rwandan violence and high death rates.
2. Probability of death is no longer reduced by a factor of 10, as simulation during a crisis period requires a shorter time frame.
3. Higher local density of perpetrators is allowed to increase the probability of death. This was noted as a limitation of the prior version of the model and is adjusted here.
4. In order for an out-group member to be at risk of death, the *sum* of local perpetrators and complicit bystanders must exceed the number of active bystanders. The prior logic required perpetrators to be in the presence of a complicit bystander in order to potentially harm the out-group. This new logic is more realistic, as perpetrators acting alone should have a chance of success, with that probability increasing when complicit bystanders are nearby.

Algorithm 6 below shows the updated logic corresponding to points 2-4 above:

Algorithm 6 Violence Against Out-Group–Increased Death Rate (adapted from [51])

let $activeBystCt$ = *number of in-group active bystanders in local radius*
let $complicitBystCt$ = *number of in-group complicit bystanders in local radius*
let $perpetratorCt$ = *number of perpetrators in local radius*
let $totalBystandPerpCt$ = *total number of all in-group civilians and perpetrators in local radius*
if $(perpetratorCt > 0) AND ((perpetratorCt + complicitBystCt) > activeBystCt))$ **then**
  $probabilityofDeath \leftarrow (\frac{complicitBystCt + perpetratorCt - activeBystCt}{totalBystandPerpCt})$
end if

7.1.4 Summary of Experimental Settings and Methodology

The validation experiment implements the following settings:

Experimental conditions and settings:

- **Condition:** Condition 3 (AB=False, IF=True) from Table 6.1, Crisis simulation–perpetrators increase from 0 to 27 agents on day 100
- **Independent variables:** Day
- **Dependent variable:** Out-Group Count
- **Other experimental settings:**
  - Number of agents: 1000[2]
  - In-Group initial percentage: 85% (850 agents)
  - Out-Group: 150 agents
  - Number of perpetrators: see Conditions section above
  - β: experiment dependent, see methodology below
  - RadiusofSight = 10 NetLogo patches
  - ProbabilityofMutation = 1 in 10,000
  - ProbabilityofDeath = 1 in 25,000
  - SusceptibilityFraction = 10,000
  - MaximumDays = experiment dependent, see methodology below
  - Environment size = 75x75 NetLogo patches

The formal validation methodology is as follows:

Part 1–Rwanda. With the Eris2 model calibrated for the Rwandan scenario using the experimental conditions and settings above, search the β parameter space to determine a β_{Rwanda} value that yields a death rate matching the real-world event: deaths of 82% of the out-group over approximately 100 days.

Part 2–Rwanda with Côte d'Ivoire Data. Leaving the Eris2 model calibrated for the Rwandan scenario, test three values of β based on the results from the data analysis presented in Section 5.3.3 and Table 5.5 over approximately 2 years from the onset of violence (830 days total):

[2] The finding from the previous version of the model presented in Sections 6.4 and 6.5 showed that results were inconsistent for different population sizes and perpetrator counts. Given the model revisions presented in 7.1.3, we have elected to run these experiments on a larger population group. To ensure that there is not a major difference in the findings for 500 agents, we tested the model at this population size for $\beta = 0.12$. In this case, the out-group was reduced by 82% in 105 days, with a standard deviation of 19. This is slightly higher than the result presented in Table 7.1; however, it is within the range of the standard deviation. Future work will fully explore the behavior of Eris2 across different population sizes and perpetrator densities.

- β_{Cote1} = 141% of β_{Rwanda} (minimum change, represents difference between overall means between results)
- β_{Cote2} = 280% of β_{Rwanda} (median change, represents median difference between mean result for Côte d'Ivoire and Rwanda's post-invasion mean)
- β_{Cote3} = 418% of β_{Rwanda} (maximum change, represents full difference between mean result for Côte d'Ivoire and Rwanda's post-invasion mean)

This experiment determines if the empirical findings of ecosystem orientation in the case of Côte d'Ivoire, relative to the finding for Rwanda, is sufficient to restrain genocide even in the extreme circumstances of the latter scenario.

Part 3–Negation Experiment. With the lowest β_{Cote} value that yields limited violence over 2 years from the onset of violence (830 days total), determine the number of perpetrators required to negate the protection that level of societal restraint offers.

Note that the results for the above experiments are taken over 25 replicates. Determining the stability of these results over 100 replicates or more is left for future work.

7.2 Experimental Set 2 – Results

Part 1–Rwandan Scenario Results. After calibrating the model according to the settings and conditions listed in Section 7.1.4, we found the following:

Table 7.1: Results for β value testing on Rwandan scenario. Experiments determine approximate value that reproduces the Rwandan death rate in which 82% of the Tutsi population (out-group) was killed in 100 days.

β value	Average days after start of crisis until 18% of out-group remains	Standard deviation over 25 replicates
0.13	110.08	16.01
0.125	107.16	18.11
0.12	100.64	16.81

It is important to note that these are preliminary findings intended to work towards proof-of-concept at this stage of the research. Future work will involve a more granular exploration of the β parameter space, with a minimum of 100 replicates for each. Given the above, we take $\beta_{Rwanda} = \mathbf{0.12}$ as the baseline for the Rwandan scenario.

Part 2–Rwandan Scenario with Ivorian Data Results. Section 7.1.4 outlined the corresponding β values for three levels of difference based on the results of the sentiment analysis presented in Section 5.3.3. Table 7.2 below details the results with $\beta_{Rwanda} = 0.12$.

Table 7.2: Results for Rwanda scenario tested with β values adjusted to reflect ecosystem orientation from Côte d'Ivoire. Experiments determine which level of increase is required to avoid genocide.

β value	Average out-group deaths in 830 days	Standard deviation over 25 replicates	Percentage of total out-group
$\beta_{Cote1} = 0.29$	119	10	80%
$\beta_{Cote2} = 0.46$	11	4	7%
$\beta_{Cote3} = 0.62$	0.5	0.5	0.3%

The results show that revising the β upward to account for the percentage ecosystem increase from the Rwandan mean over all data and the Ivorian mean, β_{Cote1}, is insufficient to prevent a genocide. However, recall from Section 5.3.3 that following the start of Tutsi-led invasions from Uganda several years before the 1994 genocide, the ecosystem mean from the Rwandan speeches drops dramatically. Because we cannot currently estimate the degree of latency presidential or elite discourse has in a society, we took β_{Cote2} as a midpoint between the minimum and maximum possible difference in means between the results. It is clear that there are strong restraints on violence for this value, with deaths of only 7% of the out-group over two years. However, the number of civilian deaths in Côte d'Ivoire during its first civil war was far lower, with 1,000 deaths in 2003, a small fraction of the total population of any of its ethnic groups [31]. β_{Cote3} represents the full magnitude of difference between the lowest level of Rwandan post-invasion mean, and introduces sufficient restraints to prevent violence and genocide over a two year time span. While this result is encouraging, suggesting that conditions in Côte d'Ivoire society were influential in preventing a genocide, there is much future research required to establish how long one can expect the tone of a president's dialogue to influence the population. In the Rwandan case, is it sufficient to use the lower mean value from three years before the genocide, or does the higher ecosystem orientation of the prior years continue to have an influence?

Figure 7.1 below provides an additional perspective on what is happening in each of these scenarios. Here, we can confirm the finding presented in Section 6.3.2 and Figure 6.6: active bystanders do not need to be a majority in order to fully protect the out-group. Even with a mean increased by 418%, the number of active bystanders is approximately half that of the passive; however, with a population size that is more than double that of the baseline case, they can effectively protect the out-group.

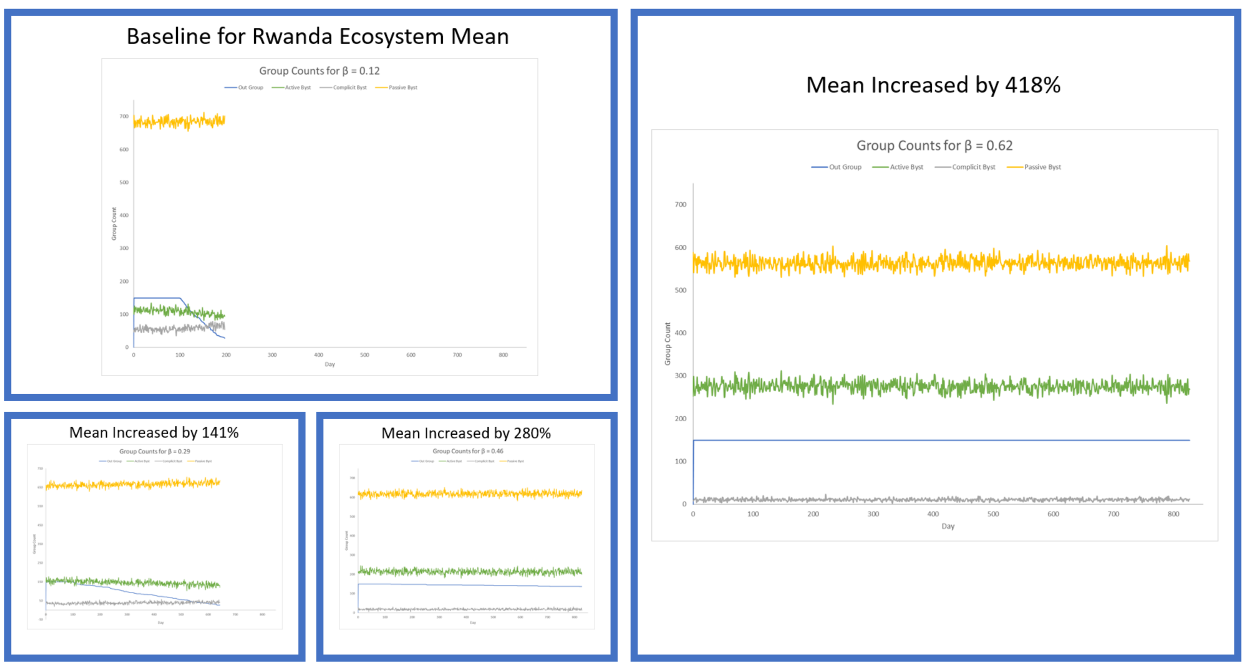


Figure 7.1: Comparison of bystander group counts for experiments increasing Rwandan β according to Ivorian data.

Part 3–Negation Experiment Results. Table 7.3 presents the results of the final experiment. In brief, these results show that if the environment contains approximately 214 to 215 perpetrators, $\beta_{Cote3} = 0.62$ is no longer sufficient to prevent a genocide. Note that this number of perpetrators is equivalent to about 25% of the in-group population. While this number is much higher than one would expect in a real-world scenario, it does show that there exists a threshold beyond which factors of restraint as modeled here can no longer protect the out-group from violence.

Table 7.3: Results from experiment negating prevention of genocide with $\beta_{Cote3} = 0.62$ by increasing perpetrator count.

Number of perpetrators	Average days after start of crisis until 18% of out-group remains	Standard deviation over 25 replicates
200	130	17
210	113	13
214	100	13
215	99	14

7.3 Experimental Set 2 – Analysis Summary

The following are the major findings from this set of experiments:

- The Eris2 model can be calibrated to reproduce general patterns of a real-world scenario. In this experimental set, the model reproduced the approximate death rate of the 1994 Rwandan genocide.
- After determining the baseline β value aligning with the Rwandan death rate, the Eris2 model was capable of modeling a relative adjustment of β based on the mean ecosystem findings for the Ivorian case presented in Section 5.3.3 and Table 5.5. The model showed that while the smallest gap in mean values was insufficient to prevent genocide, violence was greatly decreased for a median value, and nearly eliminated for the maximum possible difference in means between the two scenarios.
- The model showed that with a perpetrator count that is approximately 25% of the size of the out-group, the previous finding is negated, and genocide occurs with the original death rate.

As discussed in part in the sections above, this work has a number of limitations. The results presented here should be taken as proof of concept regarding the general capabilities of the combination of the model framework, data, and methodology to study the problem of genocide. In order to properly verify and validate the Eris2 model, it must be tested across its parameter space, similar to what is presented in Chapter 6.

7.4 Contributions, and Conclusions

The above results confirm the hypothesis presented in Section 7.1. While the smallest difference in means between the Rwandan and Ivorian data was not sufficient to prevent a genocide from occurring in the Rwandan scenario, other measures taken from the results presented in Section 5.3.3 were sufficient to ensure out-group protection in this case. This result provides a partial answer to Research Question 3 (Section 1.2.3), which asks if the model can be useful to researchers from a range of disciplines, and if it can be validated against real-world scenarios. While the results of Chapter 6 showed the Eris model's usefulness for exploring general dynamics of the problem, the Eris2 model presented here can reproduce the death rate of the 1994 Rwandan genocide, as well as employ quantitative measures from a negative case of genocide to evaluate the effectiveness of these restraints in a different scenario.

Finally, returning to the research gap outlined in Section 3.4.3, generalized models are not easily calibrated to a specific scenario, and scenario specific models cannot be easily generalized, the work presented in this chapter shows the Eris2 model's ability to provide an alternative simulation environment. The focus of this portion of the research was to connect system-level factors of restraint from different scenarios through the model. While we have yet to use this model to simulate two discrete events, it was sufficiently easy to calibrate it to the Rwandan genocide. We expect to be able to confirm the ability of Eris2 to simulate additional scenarios, and use it to perform counterfactual analysis for the Rwandan genocide and others.

Contributions. To conclude, the Eris2 model contributes the following to the field of computational social science:

1. Eris2 is a generalized agent-based model of identity-based conflict with the potential

for genocide that can be easily calibrated to model high-level dynamics of specific scenarios.

2. The model provides an environment in which the user can explore different outcomes given data measuring societal-level factors of restraint, with one example of that data being ecosystem measures as presented in Chapter 5.
3. The combination of the Eris2 model and the results from sentiment analysis of elite speech presented in Chapter 5, is a framework for better understanding the problem of genocide. It provides a synthetic environment in which to explore the effect of restraining factors in a society, and can be used to compare and contrast dynamic differences between positive and negative cases of genocide.

There are many research questions that remain to be answered, and related hypotheses to confirm or disprove. The research and findings presented in this dissertation provide a framework from which to build a robust research agenda. Chapter 8 summarizes the findings, limitations, and contributions of this work. It then looks to the future by presenting a potential structure for that research agenda, with an overarching goal of using computing technology to contribute to a more peaceful world.

CHAPTER 8: CONCLUSIONS

The research presented in this dissertation was motivated by the desire to create a novel, agent-based model and framework for understanding the dynamics of genocide. Achieving this goal required close interdisciplinary collaboration between computer scientists, historians, and social scientists. The result is a research model and framework that are the foundation for extensive future work. This chapter provides a brief review of the research, key findings, limitations, and contributions (Section 8.1), a discussion of higher-level lessons and takeaways from the work as a whole (Section 8.2), a proposed agenda for future research (Section 8.3), and concluding thoughts (Section 8.4).

8.1 Summary of the Research, Key Findings, Limitations, and Contributions

The main product of this research is the Eris model, which is grounded in political science theory of escalating and restraining factors as detailed in Section 2.2, social psychology theories of bystanders and their motivations as they relate to the need for self- and species preservation presented in Section 2.3, and complexity theory as a framework through which to combine these elements as explained in Sections 3.1, 3.3, and 3.5.

8.1.1 Review of Model Framework, Data, and Experimental Approach

The model described in Chapter 4 is fairly simple, addressing Research Question 1 (Section 1.2.1): what is an optimal, small and efficient model that can simulate conflict that may lead to genocide, and provide insight into the dynamics of the problem? The Eris environment is a world with two identity groups, an in-group and an out-group. In this world, there can be different levels of restraining factors that prevent violence against the out-group by increasing the likelihood that in-group agents will become active bystanders. These factors of restraint are captured in the model's global β function. In-group agents have attributes of identity, influence, susceptibility, egosystem, and ecosystem. Egosystem and ecosystem attributes quantify an agent's motivational orientation as related to its need for self- and species-preservation respectively. Depending on the balance of its motivational states, an in-group agent decides which type of bystander it will become. Higher egosystem leads to a greater probability of becoming a complicit bystander. Higher ecosystem leads to a greater probability of becoming an active bystander. If an agent chooses neither complicit nor active, it will be a passive bystander. Finally, perpetrators move through the world attempting to harm the out-group. Higher ratios of active bystanders create a lower likelihood that a perpetrator will achieve its goals.

The Eris model was designed to address the research gap noted in Section 3.4.3, that generalized models of conflict can not easily made specific, and the specific cannot be easily generalized. To address this gap, the model remains very general through the use of a basic theory of motivation (see Section 2.3.2); however, the β function allows it to simulate scenario specific factors of restraint in a society. This element, combined with the ability to calibrate agent and perpetrator population sizes and ratios, make Eris a flexible framework. This leads to Research Question 2 (Section 1.2.2): what is the most useful data for informing and empirically validating this model?

Chapter 5 detailed the selected data and its analysis. Given the model's reliance on agent adaptation with respect to individual motivations in the presence of global factors of restraint, this part of the work sought to inform the global β function by analyzing the emotional content of presidential speeches from a positive and negative case of genocide as provided by Scott Straus [12]. By developing a composite emotional score to quantify the ecosystem motivation in the text of these presidential speeches for Rwanda and Côte d'Ivoire as positive and negative cases respectively, this work yielded a quantitative measure that can inform the Eris model, capturing real-world dynamics for different scenarios.

The first set of explorations with the Eris model were in a generalized scenario. Chapter 6 detailed the methodology, results, and analysis of these experiments. The next step was to move toward empirical validation, which required model modifications. Validation experiments with Eris2 were covered in Chapter 7. These were the first steps in effectively combining the results of the presidential speech analysis with the agent-based model framework in order to simulate the emergence and non-emergence of genocide based on historic events. All of the above worked toward answering Research Question 3 (Section 1.2.3): is the model useful to researchers studying this problem, providing understandable and valid results? Additionally, is it flexible between scenarios to allow for analysis of different cases? The key findings below summarize the answers to these, and the other research questions stated in Section 1.2.

8.1.2 Summary of Key Findings

The following are the major findings of this research:

- Chapter 5: The composite measure of ecosystem as defined in Equation 5.1 yielded quantifiable differences between the emotional content of the speeches from Rwanda and Côte d'Ivoire [12].
- Section 6.3: System-level factors of restraint, modeled with the β function, have an exponentially beneficial effect on out-group death rates. This confirms the finding in [51]. As such, policy initiatives that can move a society toward inclusion can have exponentially beneficial effects for out-group members under certain circumstances.
- Section 6.3: Active bystanders do not need to be a majority in the environment in order to provide effective protection to the out-group. For sufficiently high β values, passive bystanders can remain a majority, and the number of active bystanders will still be enough to prevent high death rates.
- Sections 6.6 and 6.7: The Eris model can be adjusted to simulate a crisis trigger point in a society (6.6), as well as changing global levels of restraint against out-group violence (6.7).
- Section 7.1.3: The Eris model required revisions in order to properly reproduce scenarios with a high death rate. Eris2 achieves this by modifying the logic determining

the probability of an out-group member's death to be higher by increasing the effect size, and allowing for higher local perpetrator density to create more risk.

- Section 7.2: The Eris2 model can be calibrated to reproduce general patterns of a real-world scenario. The model reproduced the approximate death rate of the 1994 Rwandan genocide.
- Section 7.2: After determining the baseline β value aligning with the Rwandan death rate, the Eris2 model was capable of modeling a relative adjustment of β based on the mean ecosystem findings for the Ivorian case presented in Section 5.3.3 and Table 5.5. The model showed that while the smallest gap in mean values was insufficient to prevent genocide, violence was greatly decreased for a median value, and nearly eliminated for the maximum possible difference in means between the two scenarios. The effect can be negated by increasing the perpetrator count to be approximately 25% of the out-group size.

8.1.3 Summary of Limitations

Computational models are by their nature abstract. As such, there are significant limitations to this work due to the assumptions made in developing the model and analyzing corresponding data. Many limitations and assumptions are mentioned throughout Chapters 4, 5, 6, and 7, and the following is a brief summary of the most significant of these:

- Chapter 4: Identity is considered salient and is not multi-dimensional for agents in the model. This limitation prevents the current version of the model from capturing real-world complexities of social identity.
- Chapter 4: The out-group does not adapt in the Eris model. These agents are helpless in the face of perpetrator violence, with only in-group active bystanders providing protection. This is a significant simplification of the out-group, as it is likely that persecuted people will act to protect themselves and their communities.
- Chapter 4: In-group civilian cannot become perpetrators, and perpetrators cannot leave that role in the model. A more realistic framework would allow a threshold to

determine when an in-group civilian may choose to be a perpetrator, and vice-versa.

- Chapter 4: It is assumed that an in-group civilian with a high egosystem value will act in the interest of its own group when becoming a complicit bystander. Motivations working toward "self-preservation" are taken to be in harmony with preservation of the individual's identity group.
- Chapter 5: Analysis was completed on machine translated text, using a modern English language lexicon. Additionally, the lexicon does not include scenario specific terms that could be recognized as hate speech. While this limitation did not prevent the sentiment analysis approach from yielding measurable differences between the data, future work should strive to find a methodology that works more closely with the original text and the language of its era.
- Chapter 5: While the comparison of ecosystem scores between the Rwandan and Ivorian cases was extremely useful, they only have meaning in this context. There is not currently a baseline that indicates what type of speech represents the lowest or highest scores, limiting the usefulness of the absolute values of this measure.
- Chapter 6: The ability of the Eris model version used for experiments in this chapter was limited in terms of its ability to simulate the dynamics across different population sizes. This limitation was the result of logic that did not increase the probability of out-group death given a higher density of perpetrators; rather, this probability only increased according to the number and ratio of complicit bystanders.
- Chapter 7: The results presented in this chapter are only proof of concept, showing that the Eris2 model can be calibrated to reproduce the Rwandan death rate, and that the β can be informed by data from a negative case of genocide to show the effectiveness of factors of restraint. Future work should more carefully calibrate the model to the exact ratios of identity groups and perpetrators, and also incorporate death and reproduction rates that correspond to the scenario.

8.1.4 Summary of Contributions

The following are the major contributions of this research:

- The model moves social science research in the area from qualitative to quantitative in a novel way.
- The Eris model framework is an artificial world in which to explore the difficult problem of genocide. It achieves this by taking an approach informed by complexity theory and complex adaptive systems, which results in a simple model that provides useful quantitative measures.
- Eris is a generalized agent-based model of identity-based conflict and genocide that captures commonalities across societies such that it can be easily calibrated for different real-world scenarios.
- The ecosystem and egosystem framework allow this agent-based model to operationalize the behavior of people. This novel connection between the fields of natural language processing and computational social science allows for empirical validation across cases.

8.2 The Big Picture – Where To From Here?

The above section summarized the results of implementing and exploring the Eris model. While the section provides a concise overview of the work and its results, it does not fully address the potential impact of this line of research. Returning to our discussion of complexity theory and computational social science (CSS) as they inform the Eris model's theoretical framework in Chapter 3, the bottom-up approach of an agent-based model has a unique potential to inform policy work as well as social science research.

With respect to policy initiatives, when complexity theory informs CSS, the insights and answers it yields can improve understanding of societal dynamics, leading to more effective and beneficial policies. In general, this move is away from "top-down" solutions that do not account for the complexity of human interaction and adaptation, and toward more holistic policy initiatives that do. Squazzoni explores how complexity theory can inform social policies, finding that standard top-down approaches are often ineffective or counter-productive. He states:

> We live in a complex and largely unpredictable world where boundaries between technology, economic, and social systems are more and more porous. This implies that changes in one system, even apparently insignificant ones, might have serious consequences on another system's behavior. This in turn could have implications on the original system in a cumulative, nonlinear way [141].

Returning to the contributions of the Eris model, it is important to note that with continued advancements, this framework has the potential to inform policy. Introducing greater degrees of individuality as discussed in Section 8.3 below will allow Eris to explore how micro-level dynamics can impact emergent outcomes in a more robust manner. For example, if we account for the individual perception of a government's legitimacy, how much does that mitigate the effects of system-level factors of restraint? If policy initiatives geared toward connecting different identity groups in a positive manner are implemented in a specific neighborhood, how might that impact the society as a whole? How important are social networks in helping spread messages of inclusiveness, or divisiveness, and what are the thresholds beyond which out-group persecution cannot occur? Each of these questions looks at the system from the level of the individual, accounting for the importance of complex human interaction and adaptation in determining potential outcomes.

8.3 Future Research Agenda

This section outlines future work related to the Eris model, data and analytic approaches that can inform the model environment, and collaborations to ensure the validity and impact of the research. As for potential enhancements to the Eris model, particularly the Eris2 version given its usefulness for empirical validation, we plan to:

- Fully verify and validate Eris2, exploring its parameter space, and performing empirical validation in a more robust manner.

- Add an agent attribute to Eris2 that simulates individual perceptions of government legitimacy, similar to Epstein's model of civil violence [99]. Here the legitimacy attribute determines how likely an agent is to be influenced by global factors of restraint as communicated by the elite. Given a model environment calibrated to produce a peaceful society, variations in this attribute can inform the user of thresholds beyond which genocide cannot be prevented by governmental influence.
- Develop functionality that allows the user to specify neighborhoods of different density and identity group ratios in order to explore how policy initiatives affecting one neighborhood spread to others.
- Introduce dimensionality into the identity attribute in order to explore the influence of complex identity types on outcomes.
- Introduce an attribute accounting for salience of identity in order to explore how degrees of salience impact outcomes, particularly when there are in-group members for whom identity is not a major focus.
- Adjust the model such that in-group civilians can evolve into perpetrators, and perpetrators to civilians by leaving the sub-national ranks.
- Introduce the potential for an in-group active bystander to be considered an "enemy collaborator" by perpetrators, giving it a risk of persecution or death due to its support of the out-group.
- Introduce a global function (α) that simulates system-level factors of escalation like war, economic crisis, and so on. This will allow the model to capture changes and different relationships between escalating and restraining factors (β) in the society.
- Allow agents to have memory of their past interactions through an individual memory stack. Agents in the current model version are only influenced by their real-time interaction with other agents. This simplification of human behavior is not realistic, as people are also influenced by their memories.
- Implement social networks as an alternate approach to modeling memory. Agents may

have more formalized and lasting memory through network connections, and the rule structure of the model can be adjusted to allow for more closely connected agents to be more influential.

An essential piece of this research is the work with data and sentiment analysis to quantify motivational orientation presented in Chapter 5. The following are avenues for future research in this area:

- Digitize and analyze speeches from other cases included in Straus' database of presidential speeches [12]. Compare and contrast the results to determine if the analysis and conclusions for the Rwandan and Ivorian data remain valid.
- Expand potential data sources to include text from news outlets, particularly opinion pages, from a variety of societies and historic periods. This will allow for a wider range of scenario analysis with the Eris model.
- Perform analysis on appropriate text from societies representing extreme examples of inclusiveness and polarization in order to develop baseline measures of ecosystem that can better inform the meaning behind absolute values from other cases.
- Continue to develop the sentiment analysis approach to more accurately analyze text in its original language with lexicons appropriate for the historic period.

Next, this future research will continue to include close collaborations with historians and social scientists. Historians help ensure that the work is properly contextualized. Social scientists provide invaluable guidance regarding the model framework, ensuring the fundamental theories are implemented in a valid manner that can yield meaningful explanations for outcomes. We intend to add policy experts to the collaboration in order to better structure the approach so that it can be useful to their work.

Finally, we intend to reach out to modeling experts in the domain of computational social science in order to expand the research beyond the realm of the agent-based model. Two potential collaborations are with system dynamics modelers, and researchers working with data-driven, predictive models. A combination of modeling approaches is likely to have a greater impact than any one taken on its own, and we look forward to maximizing the potential of technology to assist in preventing atrocities such as genocide.

8.4 Final Remarks

This research explored how concepts from complexity theory can be applied in the domain of computational social science in order to explore and better understand the phenomenon of genocide. The Eris model is significant in two ways. First, it contributes to the field of computer science by providing a creative solution to the problem of translating social science theory into useful technological tools. Second, it opens new doors to researchers and policy makers working to understand the dynamics of conflict and genocide. This is the most important outcome, because with it comes the opportunity for the research to positively contribute toward a more peaceful world for all people.

Fear not your enemies, for they can only kill you.
Fear not your friends, for they can only betray you.
Fear only the indifferent, who permit the killers and betrayers
to walk safely on the earth.
Edward Yahinsky [142, p. 217]

One of the key findings of this work is that active bystanders do not need to have a majority to effectively protect members of the out-group. This is an encouraging result, as it affirms that individual people can make a difference even if they are in the minority. Yahinsky notes that the "indifferent," or the "passive" in the Eris model, are the most dangerous actors. Indeed, they are the majority of bystanders in the model environment both when there is genocide *and* when there is peace. The model shows the importance of choosing to take an active role in protecting others who face persecution and harm. It is the combined effect of individual decisions that determine whether or not inclusiveness and peace will prevail in a polarized world.

REFERENCES

[1] greekmythology.com, “Eris.” [Online]. Available: https://www.greekmythology.com/Other_Gods/Eris/eris.html

[2] J. Eligon, “Hate Crimes Increase for the Third Consecutive Year, F.B.I. Reports,” 2018. [Online]. Available: https://www.nytimes.com/2018/11/13/us/hate-crimes-fbi-2017.html

[3] L. Doucet, “Iraq Yazidis: The ’forgotten’ people of an unforgettable story,” 2018. [Online]. Available: https://www.bbc.com/news/world-middle-east-45406232

[4] H. Ellis-Peterson, “Myanmar’s military accused of genocide in damning UN report,” 2018. [Online]. Available: https://www.theguardian.com/world/2018/aug/27/myanmars-military-accused-of-genocide-by-damning-un-report

[5] J. M. Cox, *To kill a people : genocide in the twentieth century.* New York, New York: Oxford University Press, 2017.

[6] United Nations, “Convention on the Prevention and Punishment of the Crime of Genocide,” United Nations, Tech. Rep., 1948. [Online]. Available: https://www.un.org/en/genocideprevention/documents/atrocity-crimes/Doc.1_ConventiononthePreventionandPunishmentoftheCrimeofGenocide.pdf

[7] S. Straus, “Retreating from the Brink: Theorizing Mass Violence and the Dynamics of Restraint,” *Perspectives on Politics*, vol. 10, no. 02, pp. 343–362, 2012.

[8] ——, *Making and Unmaking Nations: War, Leadership, and Genocide in Modern Africa*, 1st ed. Cornell University Press, 2015.

[9] E. Staub, “Individual and Group Identities in Genocide and Mass Killing,” in *Social Identity, Intergroup Conflict, and Conflict Reduction*, D. W. Richard Ashmore, Lee Jussim, Ed. New York, NY: Oxford University Press, 2001, ch. 7, p. 288.

[10] J. Crocker and A. Canevello, “From Egosystem to Ecosystem: Motivations of the Self in a Social World,” in *Advances in Motivation Science, Vol. 5*, 1st ed., A. Elliot, Ed. New York: Elsevier Inc., 2018, ch. 2, pp. 41–86. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2215091918300038

[11] J. M. Epstein, “Agent-Based Computational Models and Generative Social Science,” *Complexity*, vol. 4, no. 5, pp. 41–60, 5 1999.

[12] S. Straus, “African Presidential Speeches Database,” https://faculty.polisci.wisc.edu/sstraus/african-presidential-speeches-database/, 2020. [Online]. Available: https://faculty.polisci.wisc.edu/sstraus/african-presidential-speeches-database/

[13] R. Lemkin, *Axis Rule in Occupied Europe*, 1st ed. Washington, DC: Carnegie Endowment for International Peace, 1 1944. [Online]. Available: https://www.amazon.com/Axis-Occupied-Europe-Raphael-Lemkin/dp/B0000CPHJ1/ref=sr_1_2?crid=17RNXS2H1Z69X&keywords=axis+rule+in+occupied+europe&qid=1584816332&sprefix=axis+rule+in+o%2Caps%2C152&sr=8-2

[14] A. Jones, *Genocide : A Comprehensive Introduction*, 3rd ed. Abingdon, Oxon: Routledge, 2017.

[15] Merriam-Webster, "War | Definition of War by Merriam-Webster." [Online]. Available: https://www.merriam-webster.com/dictionary/war

[16] M. Shaw, *What is Genocide?*, 2nd ed. Cambridge CB2 1UR, UK: Polity Press, 2015.

[17] S. Straus, ""Destroy Them to Save Us'": Theories of Genocide and the Logics of Political Violence," *Terrorism and Political Violence*, vol. 24, no. 4, pp. 544–560, 2012.

[18] ——, *The Order of Genocide: Race, Power, and War in Rwanda*, 1st ed. Ithaca, NY: Cornell University Press, 2006.

[19] N. Robins and A. Jones, Eds., *Genocides by the Oppressed: Subaltern Genocide in Theory and Practice*, 1st ed. Bloomington: Indiana University Press, 2009.

[20] E. D. Weitz, *A Century of Genocide: Utopias of Race and Nation-Updated Edition*. Princeton University Press, 2015.

[21] J. Meierhenrich, *Genocide: A Reader*. Oxford University Press USA, 2014.

[22] D. Bloxham and A. D. Moses, *The Oxford Handbook of Genocide Studies*. Oxford University Press, 2010.

[23] B. Kiernan, *Blood and Soil: A World History of Genocide and Extermination from Sparta to Darfur*. Yale University Press, 2007.

[24] A. Weiss-Wendt, R. Krieken, A. A. Cave, B. Kiernan, D. Bergen, D. Moshman, V. Sanford, J. Docker, and R. Hitchcock, *The Historiography of Genocide*. Springer, 2008.

[25] S. Totten and W. S. Parsons, *Centuries of Genocide: Essays and Eyewitness Accounts*. Routledge, 2012.

[26] F. Chalk and K. Jonassohn, *The History and Sociology of Genocide: Analyses and Case Studies*. Yale University Press, 1990.

[27] A. L. Des Forges, T. P. Longman, J. Rone, Human Rights Watch (Organization), and Fédération internationale des ligues des droits de l'homme, *Leave none to tell the story: Genocide in Rwanda*, 1st ed. New York, NY: Human Rights Watch, 1999.

[28] United Nations, "Outreach Programme on the 1994 Genocide Against the Tutsi in Rwanda and the United Nations." [Online]. Available: https://www.un.org/en/preventgenocide/rwanda/historical-background.shtml

[29] R. Dallaire, *Shake Hands with the Devil: The Failure of Humanity in Rwanda.* Vintage Canada, 2009.

[30] "Ivory Coast profile - Timeline - BBC News." [Online]. Available: https://www.bbc.com/news/world-africa-13287585

[31] Project Ploughshares, "Armed Conflicts Report-Cote d'Ivoire (2002 - first combat deaths)," Project Ploughshares, Waterloo, Ontario, Canada, Tech. Rep., 2009. [Online]. Available: https://www.justice.gov/sites/default/files/eoir/legacy/2014/02/25/Cote_d%27Ivoire.pdf

[32] M. G. Roskin, "Political science | Britannica." [Online]. Available: https://www.britannica.com/topic/political-science

[33] E. M. von Briesen, N. G. Bacaksizlar, and M. Hadzikadic, "Modeling Genocide at the System and Agent Levels," *Journal on Policy and Complex Systems*, vol. 3, no. 2, pp. 31–48, 2017. [Online]. Available: http://www.ipsonet.org/publications/open-access/policy-and-complex-systems/journal-of-policy-and-complex-systems-volume-3-number-2-fall-2017

[34] E. Verdeja, "On genocide: Five contributing factors," *Contemporary Politics*, vol. 8, no. 1, pp. 37–54, 2002.

[35] ——, "Genocide: Clarifying Concepts and Causes of Cruelty," *The Review of Politics*, vol. 72, no. 03, pp. 513–526, 2010.

[36] J. Waller, "Perpetrators of Genocide : An Explanatory Model of Extraordinary Human Evil," *journal of hate studies*, vol. 52, pp. 5–22, 2001.

[37] J. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World.* McGraw-Hill Higher Education, 2000.

[38] S. Straus, "Triggers of Mass Atrocities," *Politics and Governance*, vol. 3, no. 3, pp. 5–15, 2015.

[39] E. Verdeja, "Predicting Genocide and Mass Atrocities," *Genocide Studies and Prevention*, vol. 9, no. 3, pp. 13–32, 2 2016. [Online]. Available: http://scholarcommons.usf.edu/gsp/vol9/iss3/5/

[40] K. Lewin, *Principles of Topological Psychology*, 1st ed. York, PA: McGraw-Hill, 1938, vol. 7, no. 1. [Online]. Available: https://archive.org/details/principlesoftopo00lewirich/page/n5/mode/2up

[41] M. Argyle, "Social psychology | Britannica." [Online]. Available: https://www.britannica.com/science/social-psychology

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

[43] E. A. Zillmer, R. P. Archer, and R. Castino, “Rorschach records of Nazi war criminals: a reanalysis using current scoring and interpretation practices.” pp. 85–99, 1989. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/2918460

[44] C. Browning, *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland*, 1st ed. New York, NY: Harper Collins, 1998.

[45] E. Staub, “Obeying, Joining, Following, Resisting, and Other Processes in the Milgram Studies, and in the Holocaust and Other Genocides: Situations, Personality, and Bystanders,” *Journal of Social Issues*, vol. 70, no. 3, pp. 501–514, 2014.

[46] G. Le Bon, *The Crowd: A Study of the Popular Mind.* Reprint, Mineola, NY: Dover Publications, 2002, 1896.

[47] H. Tajfel, “Social identity and intergroup behaviour,” *Social Science Information*, vol. 13, no. 2, pp. 65–93, 1974.

[48] W. B. Swann, J. Jetten, Ã. Gómez, H. Whitehouse, and B. Bastian, “When group membership gets personal: A theory of identity fusion.” *Psychological Review*, vol. 119, no. 3, pp. 441–456, 2012. [Online]. Available: http://doi.apa.org/getdoi.cfm?doi=10.1037/a0028589

[49] S. Roccas and M. B. Brewer, “Social Identity Complexity,” *Personality and Social Psychology Review*, vol. 6, no. 2, pp. 88–106, 5 2002. [Online]. Available: http://journals.sagepub.com/doi/10.1207/S15327957PSPR0602_01

[50] A. Levy, I. Zezelj, M. Brankovic, S. Dusanic, M. van Zomeren, T. Saguy, and E. Halperin, “Complex Social Identities and Intergroup Relations: Gateway Groups in the Western Balkans,” *Social Psychology*, vol. 50, pp. 201–206, 4 2018. [Online]. Available: https://econtent.hogrefe.com/doi/abs/10.1027/1864-9335/a000379?journalCode=zsp

[51] E. von Briesen, A. Canevello, S. Shaikh, J. Cox, and M. Hadzikadic, “Modeling Genocide: An Agent-Based Model of Bystander Motivation and Societal Restraints,” in *Proceedings of the 2019 International Conference of The Computational Social Science Society of the Americas.* Springer, 2019.

[52] E. Staub, *The Roots of Evil : The Origins of Genocide and Other Group Violence*, 1st ed. Cambridge University Press, 1989.

[53] World Bank, “Access to electricity (% of population) - Sub-Saharan Africa | Data,” 2020. [Online]. Available: https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=ZG

[54] E. Staub, *The Roots of Goodness and Resistance to Evil: Inclusive Caring, Moral Courage, Altruism Born of Suffering, Active Bystandership, and Heroism*, 1st ed. New York, NY: Oxford University Press, 2015. [Online]. Available: https://global.oup.com/academic/product/the-roots-of-goodness-and-resistance-to-evil-9780195382037

[55] A. M. Turing, "The Chemical Basis of Morphogenesis," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 237, no. 641, pp. 37–72, 1952. [Online]. Available: http://www.jstor.org/stable/92463?seq=1#page_scan_tab_contents

[56] P. Ball, "Forging patterns and making waves from biology to geology: a commentary on Turing (1952) 'The chemical basis of morphogenesis'," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 370, no. 1666, 2015. [Online]. Available: http://rstb.royalsocietypublishing.org/content/370/1666/20140218

[57] J. H. Holland, "Complex Adaptive Systems," *Daedalus*, vol. 121, no. 1, pp. 17–30, 1992. [Online]. Available: http://www.jstor.org/stable/20025416

[58] G. P. Richardson, "Core of System Dynamics," in *System Dynamics: Theory and Applications - In Encyclopedia of Complexity and Systems Science Series*, B. Dangerfield, Ed. New York, NY: Springer, 2020, pp. 11–20.

[59] C. M. Macal, "Everything you need to know about agent-based modelling and simulation," *Journal of Simulation*, vol. 10, no. 2, pp. 144–156, 5 2016. [Online]. Available: https://www.tandfonline.com/doi/full/10.1057/jos.2016.7

[60] C. Macal and M. North, "Introductory tutorial: agent-based modeling and simulation," pp. 6–20, 2014. [Online]. Available: http://dl.acm.org/citation.cfm?id=2693856

[61] S. Viridi, M. Dwitasari, M. Takaendengan, K. Sari, and U. Pasaribu, "Traffic flow simulation using agent-based model: A case of single lane with multiple traffic lights and input-output node - IOPscience," in *Journal of Physics: Conference Series, International Conference on Mathematics and Natural Sciences (ICNMS)*, West Java, Indonesia, 11 2018. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1245/1/012011/meta

[62] M. Madireddy, D. J. Medeiros, and S. Kumara, "An agent based model for evacuation traffic management," in *Proceedings - Winter Simulation Conference*, 2011, pp. 222–233.

[63] U. Wilensky, "NetLogo Models Library: Flocking," Evanston, IL, 1998. [Online]. Available: https://ccl.northwestern.edu/netlogo/models/Flocking

[64] E. Frias-Martinez, G. Williamson, and V. Frias-Martinez, "An Agent-Based Model of Epidemic Spread Using Human Mobility and Social Network Information - IEEE Conference Publication," in *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing.* Boston, MA: IEEE, 10 2011. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6113095

[65] S. H. Sabzpoushan and F. Pourhasanzade, "A new method for shrinking tumor based on microenvironmental factors: Introducing a stochastic agent-based model of avascular tumor growth," *Physica A: Statistical Mechanics and its Applications*, vol. 508, pp. 771–787, 10 2018.

[66] U. Wilensky and W. Rand, *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*, 1st ed. Cambridge, MA: The MIT Press, 2015.

[67] T. S. Alexandre, V. Costa, and R. Rego, "NetLogo Wolf-Sheep Predator-Mediated Coexistence model," Ponta Delgada., 2015. [Online]. Available: http://ccl.northwestern.edu/netlogo/models/community/Wolf-Sheep%20Predator-Mediated%20Coexistence%20(FINAL)

[68] U. Wilensky, "NetLogo Communication-T-T Network Example model," Evanston, IL, 2004. [Online]. Available: http://ccl.northwestern.edu/netlogo/models/Communication-T-TNetworkExample

[69] T. C. Schelling, "Dynamic models of segregation," *The Journal of Mathematical Sociology*, vol. 1, no. 2, pp. 143–186, 7 1971. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/0022250X.1971.9989794

[70] W. Rand and U. Wilensky, "NetLogo Models Library: Traffic Basic Adaptive," Evanston, IL, 2008. [Online]. Available: http://ccl.northwestern.edu/netlogo/models/TrafficBasicAdaptive

[71] U. Wilensky, "NetLogo," Evanston, IL, 1999. [Online]. Available: http://ccl.northwestern.edu/netlogo/

[72] T. Pike, "GitHub - projectmesa/mesa: Mesa is an agent-based modeling framework in Python," 2020. [Online]. Available: https://github.com/projectmesa/mesa

[73] C. Gunaratne, "NL4Py," Orlando, FL, 2018. [Online]. Available: https://github.com/chathika/NL4Py

[74] W. Rand and R. T. Rust, "Agent-based modeling in marketing: Guidelines for rigor," *International Journal of Research in Marketing*, vol. 28, no. 3, pp. 181–193, 9 2011.

[75] X. Zhai, D. Zhong, and Q. Luo, "Turn it around in crisis communication: An ABM approach," *Annals of Tourism Research*, vol. 79, p. 102807, 11 2019.

[76] H. Zhang and Y. Vorobeychik, "Empirically grounded agent-based models of innovation diffusion: a critical review," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 707–741, 6 2019. [Online]. Available: https://link.springer.com/article/10.1007/s10462-017-9577-z

[77] J. M. Epstein and R. Axtell, *Growing Artificial Societies: Social Science From the Bottom Up*. Brookings Institution Press, 1996.

[78] R. Conte and F. Giardini, "Towards Computational and Behavioral Social Science," *European Psychologist*, vol. 21, no. 2, pp. 131–140, 4 2016. [Online]. Available: http://econtent.hogrefe.com/doi/10.1027/1016-9040/a000257

[79] N. Gilbert and S. Bullock, “Complexity at the social science interface,” *Complexity*, vol. 19, no. 6, pp. 1–4, 7 2014. [Online]. Available: http://doi.wiley.com/10.1002/cplx.21550

[80] D. Lazer, A. Pentland, L. Adamic, S. Aral, A. L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne, “Social science: Computational social science,” *Science*, vol. 323, no. 5915, pp. 721–723, 2 2009.

[81] R. Conte and M. Paolucci, “On agent-based modeling and computational social science,” *Frontiers in Psychology*, vol. 5, no. JUL, p. 668, 7 2014.

[82] B. Harff, “No Lessons Learned from the Holocaust? Assessing Risks of Genocide and Political Mass Murder since 1955,” *American Political Science Review*, vol. 97, no. 1, pp. 57–73, 2003. [Online]. Available: http://www.jstor.org/stable/pdf/3118221.pdf

[83] E. Bruch and J. Atwell, “Agent-Based Models in Empirical Social Research,” *Sociological Methods & Research*, vol. 44, no. 2, pp. 186–221, 2015. [Online]. Available: http://journals.sagepub.com/doi/pdf/10.1177/0049124113506405

[84] “Santa Fe Institute.” [Online]. Available: https://www.santafe.edu/

[85] M. Gell-Mann, *The Quark and the Jaguar: Adventures in the Simple and the Complex.* W H Freeman & Co, 1994. [Online]. Available: http://uncc.worldcat.org/title/quark-and-the-jaguar-adventures-in-the-simple-and-the-complex/oclc/29794942&referer=brief_results

[86] R. Axtell, R. Axelrod, J. M. Epstein, and M. D. Cohen, “Aligning simulation models: A case study and results,” *Computational and Mathematical Organization Theory*, vol. 1, no. 2, pp. 123–141, 2 1996. [Online]. Available: http://link.springer.com/10.1007/BF01299065

[87] R. Bhavnani, M. G. Findley, and J. H. Kuklinski, “Rumor Dynamics in Ethnic Violence,” *The Journal of Politics*, vol. 71, no. 3, pp. 876–892, 7 2009. [Online]. Available: http://www.journals.uchicago.edu/doi/10.1017/S002238160909077X

[88] B. Harff, “Detection: The History and Politics of Early Warning,” in *Responding to Genocide: The Politics of International Action.* Lynne Rienner Publishers Boulder/London, 2013, ch. 4, pp. 85–110.

[89] J. Goldstone, “A Global Model for Forecasting Political Instability,” *American Journal of Political Science*, vol. 54, no. 1, 2010.

[90] N. Rost, “Will it happen again? On the possibility of forecasting the risk of genocide,” *Journal of Genocide Research*, vol. 15, no. 1, pp. 41–67, 3 2013. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/14623528.2012.759398

[91] B. E. Goldsmith, C. R. Butcher, D. Semenovich, and A. Sowmya, “Forecasting the onset of genocide and politicide: Annual out-of-sample forecasts ona global dataset, 1988-2003,” *Journal of Peace Research*, vol. 50, no. 4, pp. 437–452, 7 2013. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0022343313484167

[92] B. E. Goldsmith and C. Butcher, “Genocide Forecasting: Past Accuracy and New Forecasts to 2020,” *Journal of Genocide Research*, vol. 20, no. 1, pp. 90–107, 2018. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/14623528.2017.1379631

[93] C. Butcher, B. E. Goldsmith, S. Nanlohy, A. Sowmya, and D. Muchlinski, “Introducing the Targeted Mass Killing Data Set for the Study and Forecasting of Mass Atrocities,” *Journal of Conflict Resolution*, p. 002200271989640, 1 2020. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0022002719896405

[94] United Nations, “United Nations Office on Genocide Prevention and the Responsibility to Protect.” [Online]. Available: https://www.un.org/en/genocideprevention/

[95] “Genocide Watch - Ten Stages of Genocide World Map.” [Online]. Available: http://www.genocidewatch.com/ten-stages-of-genocide-world-map

[96] International Crisis Group, “Crisis Group.” [Online]. Available: https://www.crisisgroup.org/

[97] The Sentinel Project, “The Sentinel Project |.” [Online]. Available: https://thesentinelproject.org/

[98] Global Centre for R2P, “Home - Global Centre for the Responsibility to Protect.” [Online]. Available: https://www.globalr2p.org/

[99] J. M. Epstein, “Modeling civil violence: an agent-based computational approach.” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 3, pp. 7243–7250, 2002.

[100] R. A. Hammond and R. Axelrod, “The Evolution of Ethnocentrism,” *Journal of Conflict Resolution*, vol. 50, no. 6, pp. 926–936, 12 2006. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0022002706293470

[101] R. Bhavnani, “Ethnic Norms and Interethnic Violence: Accounting for Mass Participation in the Rwandan Genocide,” *Journal of Peace Research*, vol. 43, no. 6, pp. 651–669, 2006.

[102] R. Bhavnani and H. J. Choi, “Modeling civil violence in Afghanistan: Ethnic geography, control, and collaboration,” *Complexity*, vol. 17, no. 6, pp. 42–51, 7 2012. [Online]. Available: http://doi.wiley.com/10.1002/cplx.21399

[103] N. B. Weidmann and I. Salehyan, “Violence and Ethnic Segregation: A Computational Model Applied to Baghdad,” *International Studies Quarterly*, vol. 57, no. 1, pp. 52–64, 3 2013. [Online]. Available: https://academic.oup.com/isq/article-lookup/doi/10.1111/isqu.12059

[104] R. Bhavnani, K. Donnay, D. Miodownik, M. Mor, and D. Helbing, "Group Segregation and Urban Violence," *American Journal of Political Science*, vol. 58, no. 1, pp. 226–245, 1 2014. [Online]. Available: http://doi.wiley.com/10.1111/ajps.12045

[105] R. Bhavnani and D. Miodownik, "Ethnic Polarization, Ethnic Salience, and Civil War," *Journal of Conflict Resolution*, vol. 53, no. 1, pp. 30–49, 2 2009. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0022002708325945

[106] N. B. Weidmann, "Micro-cleavages and violence in civil wars: A computational assessment," *Conflict Management and Peace Science*, vol. 33, no. 5, pp. 539–558, 11 2016. [Online]. Available: http://cmp.sagepub.com/cgi/doi/10.1177/0738894215570433

[107] A. Kustov, "How ethnic structure affects civil conflict: A model of endogenous grievance," *Conflict Management and Peace Science*, vol. 34, no. 6, pp. 660–679, 11 2017. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0738894215613035

[108] Y. Ibrahim and R. Hassan, "A revolutionary crowd model: Implemented to contrast oscillating to consistent media influence on crowd behavior," *SIMULATION*, vol. 93, no. 11, pp. 951–971, 11 2017. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0037549717702722

[109] F. L. Shults, R. Gore, W. J. Wildman, C. Lynch, J. E. Lane, and M. Toft, "A Generative Model of the Mutual Escalation of Anxiety Between Religious Groups," *Journal of Artificial Societies and Social Simulation*, vol. 21, no. 4, p. 7, 10 2018. [Online]. Available: http://jasss.soc.surrey.ac.uk/21/4/7.html

[110] Editors of Encyclopaedia Britannica, "Eris | Greek and Roman mythology | Britannica." [Online]. Available: https://www.britannica.com/topic/Eris-Greek-and-Roman-mythology

[111] greekmythology.com, "Harmonia." [Online]. Available: https://www.greekmythology.com/Other_Gods/Harmonia/harmonia.html

[112] J. Bednar, A. Bramson, A. Jones-Rooy, and S. Page, "Emergent cultural signatures and persistent diversity: A model of conformity and consistency," *Rationality and Society*, vol. 22, no. 4, pp. 407–444, 11 2010.

[113] S. F. Railsback, S. L. Lytinen, and S. K. Jackson, "Agent-based Simulation Platforms: Review and Development Recommendations," *Simulation*, vol. 82, no. 9, pp. 609–623, 2006. [Online]. Available: http://sim.sagepub.com/content/82/9/609.abstract

[114] R. Grider, "Optimizing NetLogo Runs · NetLogo/NetLogo Wiki · GitHub." [Online]. Available: https://github.com/NetLogo/NetLogo/wiki/Optimizing-NetLogo-Runs

[115] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, "Current State of Text Sentiment Analysis from Opinion to Emotion Mining," *ACM Computing Surveys*, vol. 50, no. 2,

pp. 1–33, 5 2017. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3071073.3057270

[116] L. Canales and P. Martínez-Barco, "Emotion Detection from text: A Survey," in *Canales, L., & MProceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC)*, 2014, pp. 37–43. [Online]. Available: http://www.scopus.com/

[117] C. Strapparava and A. Valitutti, "WordNet-Affect: an Affective Extension of WordNet," in *4th International Conference on Language Resources and Evaluation*, 2004, pp. 1083–1086. [Online]. Available: http://openmind.media.mit.edu

[118] S. Mohamad, "NRC Emotion Lexicon." [Online]. Available: http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

[119] J. Hellrich, S. Buechel, and U. Hahn, "Modeling Word Emotion in Historical Language: Quantity Beats Supposed Stability in Seed Word Selection," in *Proc. of the 3rd Joint SIGHU Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature.* Association for Computational Linguistics (ACL), 6 2019, pp. 1–11. [Online]. Available: http://arxiv.org/abs/1806.08115

[120] A. Abdaoui, J. Azé, S. Bringay, and P. Poncelet, "FEEL: a French Expanded Emotion Lexicon," *Language Resources and Evaluation*, vol. 51, no. 3, pp. 833–855, 9 2017. [Online]. Available: https://link.springer.com/article/10.1007/s10579-016-9364-5

[121] Y. Chen and S. Skiena, "Building Sentiment Lexicons for All Major Languages," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 2014. [Online]. Available: https://www.researchgate.net/publication/270877514_Building_Sentiment_Lexicons_for_All_Major_Languages

[122] M. Araujo, A. Pereira, J. Reis, and F. Benevenuto, "An evaluation of machine translation for multilingual sentence-level sentiment analysis," in *Proceedings of the ACM Symposium on Applied Computing*, vol. 04-08-April-2016. Association for Computing Machinery, 4 2016, pp. 1140–1145.

[123] N. Dennis Gitari, Z. Zuping, H. Damien, and J. Long, "A Lexicon-based Approach for Hate Speech Detection," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 10, no. 4, pp. 215–230, 2015. [Online]. Available: http://dx.

[124] R. Martins, M. Gomes, J. J. Almeida, P. Novais, and P. Henriques, "Hate speech classification in social media using emotional analysis," in *Proceedings - 2018 Brazilian Conference on Intelligent Systems, BRACIS 2018.* Institute of Electrical and Electronics Engineers Inc., 12 2018, pp. 61–66.

[125] M. Wiegand, J. Ruppenhofer, A. Schmidt, and C. Greenberg, "Inducing a Lexicon of Abusive Words: a Feature-Based Approach," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2018, pp. 1046–1056. [Online]. Available: www.urbandictionary.com

[126] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word-emotion association lexicon," in *Computational Intelligence*, vol. 29, no. 3. John Wiley & Sons, Ltd, 8 2013, pp. 436–465.

[127] Google, "googletrans 2.4.0." [Online]. Available: https://pypi.org/project/googletrans/

[128] A. Twagilimana, *Historical Dictionary of Rwanda - Aimable Twagilimana - Google Books*, 2nd ed. London: Rwoman \& Littlefield, 2016. [Online]. Available: https://books.google.com/books/about/Historical_Dictionary_of_Rwanda.html?id=9E6xCgAAQBAJ

[129] S. Straus, "Rwanda Speeches," https://faculty.polisci.wisc.edu/sstraus/african-presidential-speeches-database/rwanda-speeches/. [Online]. Available: https://faculty.polisci.wisc.edu/sstraus/african-presidential-speeches-database/rwanda-speeches/

[130] "Rwanda profile - Timeline - BBC News." [Online]. Available: https://www.bbc.com/news/world-africa-14093322

[131] Trial Watch, "Leon Mugesera - TRIAL International." [Online]. Available: https://trialinternational.org/latest-post/leon-mugesera/

[132] L. Mugesera, "Translated 1992 speech of Leon Mugesera." [Online]. Available: http://www.rwandagateway.org/spip.php?article1340

[133] Editors of Encylopaedia Britannica, "Félix Houphouët-Boigny | president of Côte d'Ivoire | Britannica," 2020. [Online]. Available: https://www.britannica.com/biography/Felix-Houphouet-Boigny

[134] S. Straus, "Cote d'Ivoire Speeches | Scott Straus." [Online]. Available: https://faculty.polisci.wisc.edu/sstraus/african-presidential-speeches-database/cote-divoire-speeches/

[135] ABBYY, "FineReader OCR Pro for Mac," https://www.abbyy.com/en-us/finereader/pro-for-mac/, 2013. [Online]. Available: https://www.abbyy.com/en-us/finereader/pro-for-mac/

[136] Radio Rwanda, "RWANDAN PRESIDENT BRIEFS CENTRAL COMMITTEE ON POLITI-CAL AND MILITARY SITUATION," BBC Summary of World Broadcasts, Tech. Rep., 12 1990.

[137] S. Nsanzimana, "Prime Minister Briefs Diplomats on Distrubances," pp. 1–2, 3 1992. [Online]. Available: https://faculty.polisci.wisc.edu/sstraus/wp-content/uploads/2015/03/11-Mar-1992-Rwanda.pdf

[138] F. Houphouët-Boigny, "Le message du Presidént," 1 1978. [Online]. Available: https://faculty.polisci.wisc.edu/sstraus/wp-content/uploads/2015/03/2-Jan-1978-Cote-dIvoire.pdf

[139] ——, “Le Presidént Houphouet-Boigny Dans Son Message De Fin D’Année: Demeurons au service du pays,” 1 1987. [Online]. Available: https://faculty.polisci.wisc.edu/sstraus/wp-content/uploads/2015/03/2-Jan-1987-Cote-dIvoire.pdf

[140] S. Straus, “How many perpetrators were there in the Rwandan genocide? An estimate,” *Journal of Genocide Research*, vol. 6, no. 1, pp. 85–98, 3 2004.

[141] F. Squazzoni, “A social science-inspired complexity policy: Beyond the mantra of incentivization,” *Complexity*, vol. 19, no. 6, pp. 5–13, 7 2014. [Online]. Available: http://doi.wiley.com/10.1002/cplx.21520

[142] H. Lookstein, *Were We Our Brothers’ Keepers: The Public Response of American Jews to the Holocaust 1938-1944*, 1st ed. New York: Hartmore House, 1985.