A GENERAL SOCIAL AGENT-BASED MODEL OF OPINION DYNAMICS WITH APPLICATIONS TO STEM EDUCATION AND RADICALIZATION

by

Joseph F. Harrison
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
In Partial fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Computational Social Science

Committee:

______________________	Dr. Claudio Cioffi-Revilla, Dissertation Director
______________________	Dr. Arie Kruglanksi, Committee Member
______________________	Dr. Ken de Jong, Committee Member
______________________	Dr. Matthew Hendrey, Committee Member
______________________	Dr. Kevin Curtin, Acting Department Chair, Dept. of Computational and Data Sciences
______________________	Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science
______________________	Dr. Peggy Agouris, Dean, College of Science

Date: ______________________

Fall 2016
George Mason University
Fairfax, VA

A General Social Agent-Based Model of Opinion Dynamics with Applications to STEM Education and Radicalization

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Joseph F. Harrison
Master of Science
George Mason University, 2009
Bachelor of Science
Utah State University, 2002

Director: Dr. Claudio Cioffi-Revilla, Professor
Department of Computational and Data Sciences

Fall 2016
George Mason University
Fairfax, VA

Dedication

For Kira.

May you never stop dancing.

Acknowledgments

I would first like to thank my advisor Claudio Cioffi-Revilla for supporting me beyond what any graduate student could expect. He provided me with guidance and opportunities without which this wouldn't have been possible. I would also like to thank the other members of my committee, Ken de Jong who provided wise counsel when I most needed it, Arie Kruglanski who shared his deep expertise on radicalization and social psychology, and Matthew Hendrey who kept me moving forward by giving his time so generously.

I am grateful to the Synergies team, John Falk, Lynn Dierking, William Penuel, Nancy Staus, and Matthew Hendrey, for envisioning and carrying out the STEM Interest project. The data they collected was incredibly valuable to me and I believe it will useful to others in my field as well. William Penuel deserves special thanks for helping to design the model rules and for providing valuable feedback on the STEM chapter.

I'm deeply indebted to Hammad Sheikh for his collaboration on the radicalization model and for sharing his expertise on data analysis, academic writing, project management, and much more. He is a brilliant scholar and a true friend.

I am grateful to Tim Gulden and Ernesto Carrella for providing perspective on the process of writing a dissertation, to Eric "Siggy" Scott and David Masad for technical assistance, to Brendon Fuhs for helping with some math, and to Peter Jones for his feedback on my presentation. I'm also grateful to Wander Jager and Frédéric Amblard for their correspondence regarding their influential work. Thank you to Jayshree Sharma for helping me with the ARGO cluster. Blessed are the tool-builders: Sean Luke for writing MASON and ECJ, and Hadley Wickham for his numerous indispensable R packages. Thanks also to Rob Axtell and Andrew Crooks for nudging me along the way, and to Karen Underwood for holding the universe together.

I want to thank my friends and family who have supported me and made this chapter of my life more enjoyable. Finally, I must express my deepest gratitude to my wife Britta for her love, support, and enduring patience.

Major funding for the Synergies project was provided by the Noyce Foundation and the Lemelson foundation. Funding for the radicalization model was provided by Minerva grant no. N00014130054 administered through the Office of Naval Research. Additional funding for this work came from MURI grant no. N00014-08-1-0921, also administered through the Office of Naval Research, and from the Center for Social Complexity at George Mason University. Large-scale experiments were run on ARGO, a research computing cluster provided by the Office of Research Computing at George Mason University, VA. (URL:http://orc.gmu.edu)

Table of Contents

List of Tables

List of Figures

Abstract

A GENERAL SOCIAL AGENT-BASED MODEL OF OPINION DYNAMICS WITH APPLICATIONS TO STEM EDUCATION AND RADICALIZATION

Joseph F. Harrison, PhD

George Mason University, 2016

Dissertation Director: Dr. Claudio Cioffi-Revilla

Many aspects of our society are affected by how opinions change and ideology spreads (e.g., interest in STEM and political radicalization), but the underlying processes are not well understood. Previous attempts at modeling these phenomena have suffered from a lack of empirical data and/or insufficient grounding in social-psychological theory. Moreover, the field of opinion dynamics would benefit from a broader view of the discipline that captures the commonalities between different domains.

This dissertation presents a general framework for agent-based modeling (ABM) of opinion dynamics called the *general model* and demonstrates it using ABMs in two significantly different domains: interest in the STEM fields (science, technology, engineering, and mathematics), and political radicalization resulting from personal grievance. Both models are novel within opinion dynamics in that they update agent opinions using rules designed in conjunction with subject-matter experts, and because they make use of domain-specific data. They each make substantial contributions in their respective areas of study.

The first model pertains to adolescents' interest in the STEM fields and utilizes rich longitudinal data gathered annually over 3 years. The model is calibrated using evolutionary computation, validated using subsequent surveys, and used to explore potential intervention strategies. Of those evaluated, knowledge brokering and increasing friend co-participation are shown to be demonstrably promising. The model is groundbreaking in many ways; it's the first ABM to model the role and effect of STEM-related activities, the first that includes the transition from situational to individual interest per Hidi and Renninger (2006), and the first to use detailed longitudinal survey data in any model of opinion dynamics.

The second domain, political radicalization, is explored using an ABM based on a psychological theory of radicalization grounded in the *Significance Quest* theory of Kruglanski et al. (2009, 2013, 2014) and the *multi-path* theory of Cioffi-Revilla (2010). The model is calibrated using data from potential jihadists in Morocco, and used to explore network effects of the psychological (i.e., individual-level) radicalization processes. It shows that the psychological processes do indeed increase the number of extremists on the group level. The model shows that when traumatic events are relatively rare, exposure to diverse opinions can reduce/prevent radicalization. This is the first ABM of radicalization based on existing social-psychological theory, the first to incorporate motivational elements, and the first to use real-word data to any significant degree. It is also the first ABM in opinion dynamics to model latitude of non-commitment directly, to use thresholds drawn from a distribution, or to dynamically modify thresholds based on exogenous events. It is also the first work in any social science to explore the intermixing of people with varying latitudes of non-commitment.

Part I

Foundations

Chapter 1: Introduction

How are opinions formed and how do they change? Are they shaped more by external events or through interactions with others? Many of the existing models of opinion dynamics have taken inspiration from different branches of physics, such as statistical mechanics (Bahr and Passerini, 1998), ising spin (Sznajd-Weron and Sznajd, 2000), among others. This dissertation instead uses lightweight cognitive mechanisms grounded in social psychology.

Human beings are complex and our cognitive processes are not always easy to discern. We must infer, from external observation, what these processes are. Gigerenzer (2007) tells the story of a baseball coach who, tired of seeing his outfielders jogging instead of sprinting, instructed them to run full-speed to the ball's destination and wait there to catch it. But when his players tried this approach, they began making more errors. The problem, it turns out, is that outfielders don't actually know where a ball is going to land when they are pursuing it. Instead, they rely on a heuristic, or rule-of-thumb, in which they begin running and adjust their speed such that the angle of their gaze remains constant, as shown in Figure 1.1.

Existing physics-based models of Opinion Dynamics generate certain phenomena (e.g., consensus formation, political polarization) fairly well, but they are overly simplistic, treating people as particles. One amusing recent example comes from (Pineda et al., 2009, p. 1) who write, "Free will is introduced in the form of noisy perturbations." This work improves on existing models by adding lightweight cognitive mechanisms based on the social psychology literature. This results in more realistic individual agent behavior and valid inferences. Another shortcoming of existing models is that they lack an empirical domain. This dissertation presents models that incorporate environmental context and real-world data.

Many lines of inquiry could be advanced by improved opinion dynamics models. On the

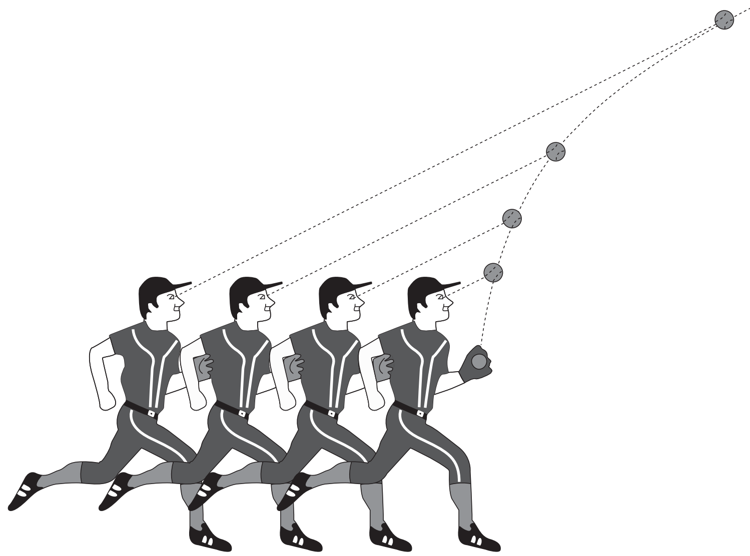


Figure 1.1: An outfielder adjusts his speed and heading such that the angle of his gaze remains constant. Source: Image from Gigerenzer (2008).

17th of December, 2010, a street merchant named Mohamed Bouazizi set himself on fire in downtown Sidi Bouzid, Tunisia to protest his unfair treatment by corrupt government officials. His self-immolation triggered a series of protests that toppled the Tunisian government, followed by the regimes in Egypt, Libya, and Yemen, and triggered major protests and political reforms in a dozen other countries across the Arab world. This "Arab Spring" caught political analysts completely by surprise. Why?

Political analysts, being human, suffer from the human tendency to expect proportionality. Push a pendulum gently and it swings backs gently, push it hard and it responds in kind. However, complex social systems can behave more like a double-pendulum; a gentle push will cause a gentle and predictable response, but a hard push will send the pendulum into violent, chaotic motion. This sort of nonlinear response is known as a critical point and often indicates a phase transition.

Improved opinion dynamics models will help analysts to better understand political tension and, perhaps, recognize when a system is nearing a critical point.

1.1 Background and Prior Literature

In the latter half of the twentieth century and beyond, social psychologists devoted much of their efforts toward understanding how a person's internal beliefs are formed and changed. This section will provide an overview of this work and describe the computational models that have been built to address the same questions. Additional review of domain-specific literature can be found for STEM interest in subsection 3.1.3, and for radicalization in subsection 4.1.3.

1.1.1 Social Psychology

Cognitive consistency

People generally tend to keep their beliefs internally consistent with one another, that is, to maintain *cognitive consistency* (Abelson et al., 1968). When a person's beliefs or actions contradict one another, it creates a sense of incongruity and discomfort known as *cognitive dissonance* (Festinger, 1962). There are many ways in which a person can reduce or eliminate cognitive dissonance. One classic study comes from Festinger and Carlsmith (1959) who had subjects perform dreary and mundane tasks for an entire hour. The subjects were randomly assigned to three groups. The first two groups were asked, after the tasks were complete, to tell the next participant (really a confederate) that the tasks were interesting. The first group was offered $1 for lying, and the second group was offered $20. The third group was used as a control group and was neither asked to lie, nor offered money. After all this was over, the subjects were interviewed by a third party (believed to be an undergraduate unrelated to the experiment) and asked if they found the tasks enjoyable, if they thought it was scientifically important, and if they would like to participate in such experiments again. The $20 group, having been paid well to lie, gave only slightly more favorable responses than the control group. The $1 group, however, gave much more positive responses to all three questions; they resolved the cognitive dissonance of lying by convincing themselves that they had actually enjoyed the tasks.

Deviating from the widespread view among theorists that cognitive consistency is a goal in itself, Kruglanski and Shteynberg (2012) argue that it is instead a means of acquiring knowledge.

Radicalization

In recent years, perhaps due to the successful prevention of large coordinated attacks, we are seeing an increase in lone wolf (or couple) attacks by homegrown terrorists. These attackers are not indoctrinated abroad in training camps run by terrorist organizations, and rarely do they have direct contact with terrorist recruiters. Instead, a process of self-radicalization takes place in which people—on their own—seek out information about radical ideologies (increasingly involving online sources and social media). There is increasing evidence from psychological research that personal grievances play an important role in this radicalization (Kruglanski et al., 2009, 2013, 2014). When people experience an acute loss of their personal significance (e.g., humiliation by others, trauma, exclusion, loss of job) they have an increased need to restore this significance. In this mental state, they become susceptible to ideologies that give their suffering meaning and provide an unambiguous way to regain the lost significance. Black and white ideologies that link the personal grievances with a bigger collective cause, name culprits (scapegoating) and provide a clear path to become a hero for the cause are especially alluring in this state. Kruglanski et al. (2014) propose that for radicalization three components have to come together: a motivational component (people are aggrieved and are motivated to regain personal significance), an ideological components (provides meaning and presents violence as appropriate means to gain significance), and a social component (group dynamics and networks which leads to the spread of the ideology and allows carrying out attacks). This model of radicalization has also been examined in the context of de-radicalization (Kruglanski et al., 2013, 2014) but that is outside the scope of this dissertation.

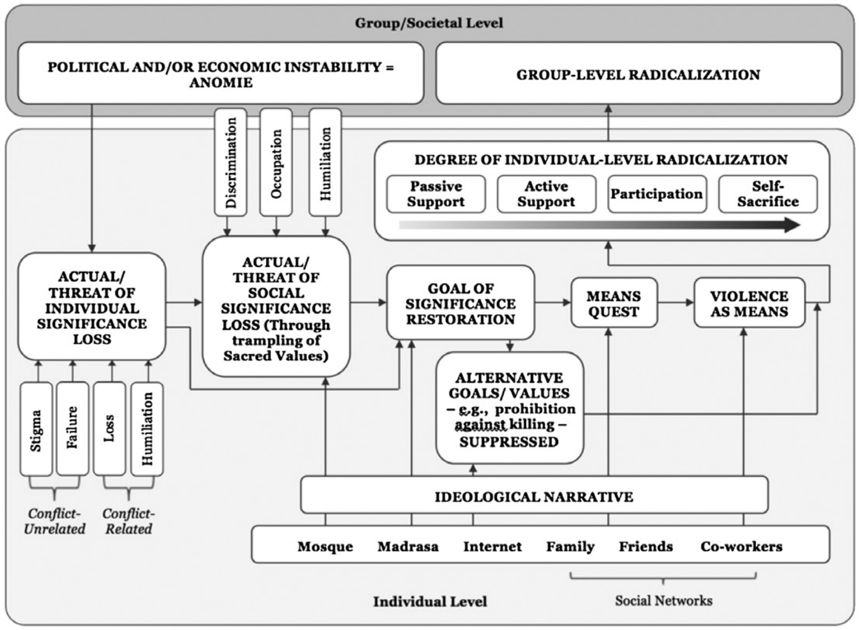


Figure 1.2: Significance Quest theory of radicalization with motivational, ideological, and social components. Source: Image from Kruglanski et al. (2014).

An illustrative example for this model is the 2015 San Bernadino attack (Schmidt and Pérez-peña, 2015) which was perpetrated by a married couple (Syed Rizwan Farook and Tashfeen Malik) and claimed 14 lives. As it turns out, the husband had difficulties at his work place and the couple had some grievances about their treatment by others. But the perpetrators seem to have situated these personal grievances within the context of the perceived global struggle of Muslims, a narrative that is at the core of jihadist ideology. This ideology replaces the uncertainties of life with meaning and certainty by blaming the grievances of Muslims on the Western societies who are portrayed as waging a "war on Islam." Beside removing ambiguity and uncertainty, it also offers a clear plan of action and path of regaining significance and becoming an instant hero: gruesome terrorist attacks. Indeed, the couple decided to act and perpetrated an attack at the work place of the husband (most victims were coworkers of his), suggesting that the personal grievance were still very much at the center of their motivations.

Dovetailing nicely with the Significance Quest (SQ) theory of radicalization (Kruglanski et al., *2009, 2013, 2014), the multi-path (MP) theory Cioffi-Revilla (2010) describes radicalization as a compound event consisting of grievance, indoctrination, and loss of the inhibition to kill (see Figure 1.3).* Together, these form a theoretical basis for the radicalization model presented in chapter 4.

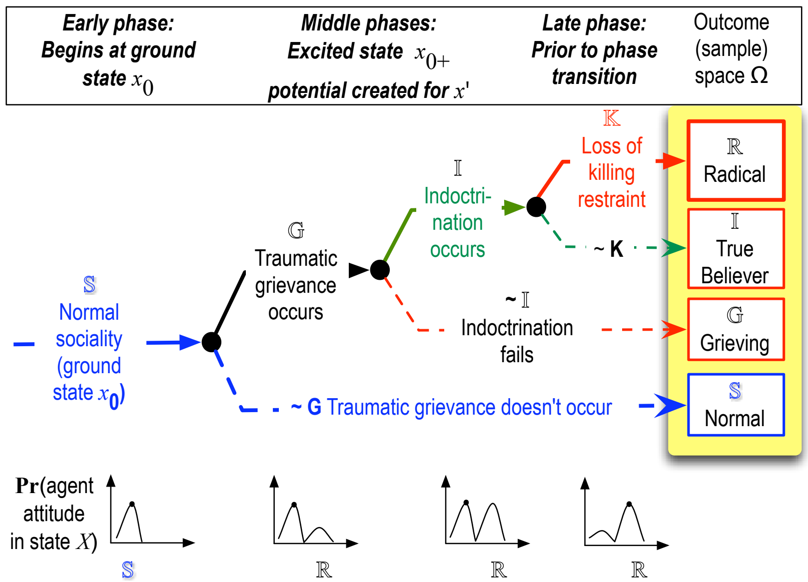


Figure 1.3: Individual radicalization as a branching process through multiple stages. Source: Image from Cioffi-Revilla (2010).

Persuasion theory In the past three decades, persuasion research in social psychology has been based largely on either the elaboration likelihood model (Petty and Cacioppo, 1986) or the heuristic systematic model (Eagly and Chaiken, 1993). The elaboration likelihood model (ELM) is based on a bounded continuum between different levels of scrutiny. At the low end, a person ignores all information relevant to the issue at hand, instead relying on peripheral cues or heuristics. At the high end, a person scrutinizes all available information and evaluates it systematically. These are referred to, respectively, as the *peripheral route* and the *central route* to persuasion. The heuristic systematic model (HSM) is based on the roughly analogous concepts of *heuristic processing* and *systematic processing*. These are also the same concepts that Kahneman (2003) describes as *system 1* and *system 2* thinking.

Indeed, the formulation of cognition as a dual process goes back at least to James (1890) and has many adherents (Chaiken and Trope, 1999; Epstein et al., 1992; Evans, 2008; Kahneman, 2003; Sloman, 1996; Strack and Deutsch, 2004). However, a growing chorus of critics (Balci, 2004; Gigerenzer and Regier, 1996; Gigerenzer and Brighton, 2009; Keren and Schul, 2009; Kruglanski et al., 2006) has highlighted the shortcomings of the dual-system approach and issued a call for a "uni-model" that combines the two systems into one coherent whole. To that end, (Kruglanski and Gigerenzer, 2011) describes some of the commonalities between intuitive and deliberative thinking. In particular, both types of judgments are based on *if-then* rules and these rules must be selected according to the cognitive constraints of the person making the judgment, and the ecological rationality of the situation.

Rules of the type "if X then Y" are generally distinguished from associations of the sort "X varies with Y" in that the former describes a directed causal link between X and Y while the latter does not. Kruglanski and Shteynberg (2012), however, argue that even associations are based on the rules, albeit of a different, bidirectional form.

Thinking deeply about an issue requires work and occurs only with sufficient motivation and cognitive resources. But what is *sufficient*? In their cognitive energetics theory (CET), Kruglanski et al. (2012) describe a competition between *driving forces* (e.g. goal importance and available cognitive resources) and *restraining forces* (e.g. task demands, resource conservation, and alternative goals). When the driving forces exceed the restraining forces, a person considers the issue carefully to reach a judgment. When the reverse is true, a person relies on heuristics or rules-of-thumb (e.g. "If the experts agree, then it's probably true"). Of course, this is not a binary determination but rather a continuous spectrum.

1.1.2 Opinion Dynamics

The field of opinion dynamics addresses how a person's opinion is affected by other people and how collective sentiments emerge from uncoordinated individual opinions. Social psychologists have been studying the impact of social pressure for many decades. In his famous conformity study, Asch (1951) showed that people will conform to a group consensus even when it is plainly wrong.

Sherif and Hovland (1961) introduced *social judgment* theory which posits that people evaluate new ideas by comparing them to their current beliefs; if similar they are accepted, if the difference is large they are rejected. Latane (1981) formalized *social impact* theory into a set of rules describing the way a person's opinion is affected by interactions with a group of people based on the strength of individual ties, the number of ties, and the salience of the event. An excellent treatment of social impact theory using agent-based modeling is provided in Gilbert and Troitzsch (2005, pp. 148–151).

This early work by social psychologists has been applied and expounded upon by political scientists, policy analysts, marketers, physicists, and researchers interested in single issues (e.g. attitudes toward climate change). In recent years, researchers have studied opinion dynamics using models implemented and analyzed computationally. The existing models can be grouped into a few broad categories.

Bounded Confidence Models

Continuous Bounded Confidence (BC) models represent an opinion as a real value from -1 to 1. For example, the opinion in question may be the person's support for a new project to widen a local highway. A value of 1 would correspond to full support of the project, while -1 indicates staunch opposition. A value of zero suggests a neutral attitude.

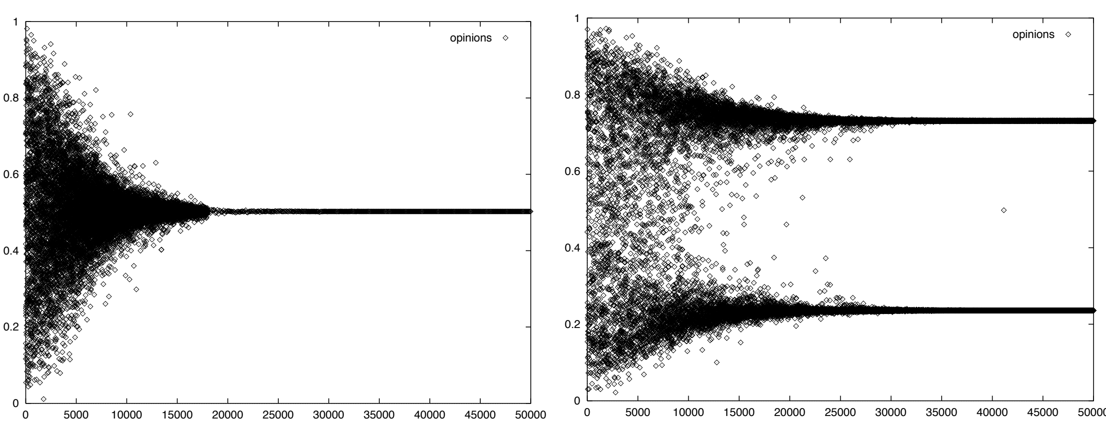
There are two main categories of BC models, those based on the work of Deffuant et al. (2000) and those based on the work of Hegselmann and Krause (2002).

Deffuant

When two agents interact in the Deffuant model, their opinions are changed in proportion to the distance between them, assuming they are within some threshold δ.[1] Consider an interaction between agents x_i and x_j, assuming their opinions are within some threshold (i.e. $|x_i - x_j| < \delta$), the result of the interaction will be:

$$\begin{aligned} x_i &= x_i + \mu(x_j - x_i) \\ x_j &= x_j + \mu(x_i - x_j), \end{aligned} \tag{1.1}$$

where μ is a tuning parameter on $[0, 1]$ that controls how much an agent's opinion is changed by any given interaction, also known as the "learning rate".



(a) $\delta = 0.5$ (b) $\delta = 0.2$

Figure 1.4: The effect of the threshold parameter in the Deffuant model with 2000 agents and a μ of 0.5. Source: Images from Deffuant et al. (2000).

The Deffuant model shows how a diverse population can reach consensus. However, a brief glance at the news makes it clear that diverse populations do not necessarily reach consensus.

[1] Deffuant's original notation used d for the threshold, which can be confused for a differential in a dynamic model. Similarly, Hegselmann used t, which can be confused for time. I'll use δ and τ respectively.

To reflect the often divisive nature of public opinion, Jager and Amblard (2005) drew upon the *social judgment* work of Sherif and Hovland (1961) and introduced a rejection mechanism by adding a second threshold. The rejection threshold t represents a point at which two opinions are so far apart, the agents are repulsed by each other and their opinions move in opposite directions. For agent i with opinion x_i, its acceptance threshold is denoted u_i and its rejection threshold is denoted t_i. When agent i interacts with agent j, its opinion is updated according to the following piecewise function:

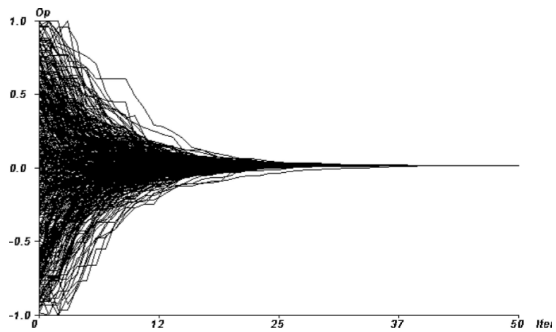
$$\begin{aligned} &\text{if}|x_i - x_j| < u_i \quad x_i = x_i + \mu(x_j - x_i) \\ &\text{if}|x_i - x_j| > t_i \quad x_i = x_i + \mu(x_i - x_j), \end{aligned} \tag{1.2}$$

where μ is the learning rate. By varying the acceptance and rejection threshold, this model can produce several different final states, as shown in Figure 1.5. With a large acceptance and large rejection threshold (1.2 and 1.6 respectively), the agents quickly converge to consensus. With a small acceptance threshold and small rejection threshold (0.4 and 0.6 respectively), the population bifurcates into two diametrically-opposed groups.

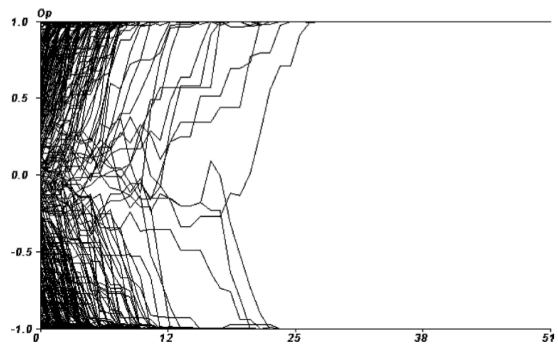
The situation becomes more interesting with a small acceptance threshold and large rejection threshold (0.6 and 1.2 respectively); the population splits into three groups with one in the middle and two at the extremes. With an even smaller acceptance threshold and larger rejection threshold (0.2 and 1.6 respectively), the model can even produce five or six opinion clusters.

Hegselmann and Krause

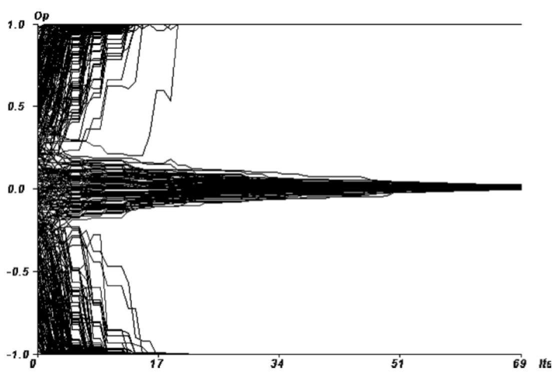
While the Deffuant formulation models interactions between a pair of agents, the Hegselmann and Krause (HK) model updates agents in groups. An agent's group contains all its neighbors that are within its acceptance threshold in opinion space. All members of the group are updated such that their opinion moves toward the group average.



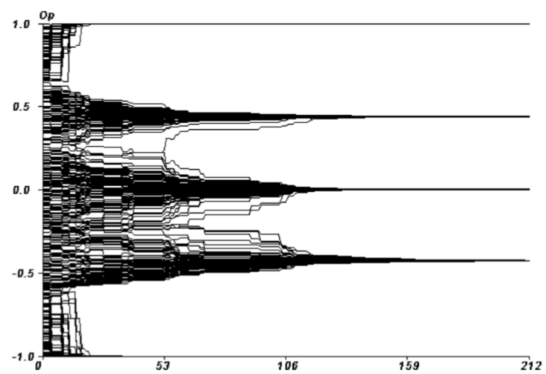
(a) U = 1.2, T=1.6



(b) U = 0.4, T=0.6



(c) U = 0.6, T=1.2



(d) U = 0.2, T=1.6

Figure 1.5: Jager and Amblard's opinion dynamics model with different values for U (acceptance threshold) and T (rejection threshold). U and T are capitalized to indicate that they are global values shared by all agents. Source: Images from Jager and Amblard (2005).

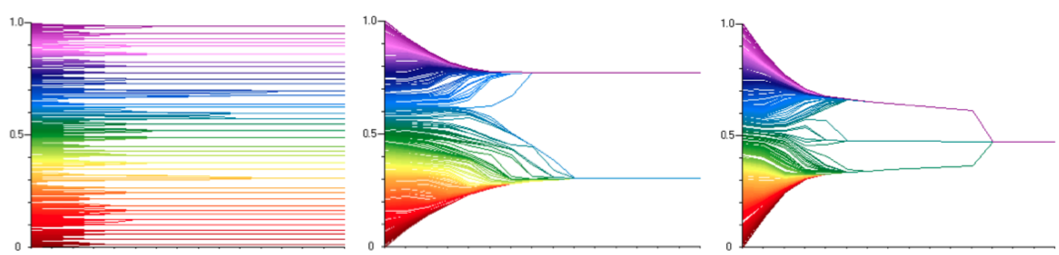
Given n agent opinions $\vec{x} = (x_1, ..., x_n)$, the set of agents with which agent i interacts is denoted $I(i, \vec{x})$ and defined as:

$$I(i, \vec{x}) = 1 \leq j \leq n \mid |x_i - x_j| \leq \varepsilon_i \tag{1.3}$$

At each step the opinions for the next step $t+1$ are updated based on the current step t as follows:

$$x_i(t+1) = |I(i, \vec{x}(t))|^{-1} \sum_{j \in I(i, \vec{x}(t))} x_j(t), \tag{1.4}$$

where $|I(i, \vec{x}(t))|$ is the number of neighbors whose opinions are within the threshold ε of agent i at timestep t. Figure 1.6 shows three runs of the HK model with different thresholds. Note how the population clusters into groups which cluster into larger groups.



(a) $\varepsilon = 0.01$ (b) $\varepsilon = 0.15$ (c) $\varepsilon = 0.25$

Figure 1.6: Hegselmann and Krause model running for 15 steps with 625 agents. Source: Images from Hegselmann and Krause (2002).

Culture Dissemination Models

Another family of opinion dynamics models is based on Robert Axelrod's cultural dissemination model (Axelrod, 1997). In his original formulation, agents have an array of integers representing their "culture." Each integer is a categorical variable representing a single cultural dimension. For example, a variable may indicate an agent's favorite sport, with 0 for baseball, 1 for basketball, 2 for football, etc. When comparing two agents on this cultural dimension, baseball (0) is not considered closer to basketball (1) than to football (2); they either match or they don't. The overall cultural distance between two agents is a count of how many non-matches they have.

Axelrod's agents are arranged in a two-dimensional lattice. When an agent is activated, it interacts with one of its neighbors, randomly chosen with a probability equal to their cultural similarity. The agent then replaces one of its cultural values with a differing value from its neighbor. After many iterations, the model eventually settles into a steady state with one or more contiguous, homogeneous groups.

Further research by Klemm et al. (2005) has shown that the model is not robust to cultural drift in the form of random perturbations. Below a certain threshold, the model results in a single homogenous group, while above the threshold, the result is a disordered, multi-cultural population.

Klemm et al. (2003) also explored interaction topologies more complex than a simple lattice, including small-world, random scale-free, and structured scale-free. They found that the final state of the model was dependent on how ordered the network was, with less-ordered networks yielding more-ordered final configurations.

The work proposed for this dissertation will not use the Axelrod model per se, however, it is a classic example of computational opinion dynamics which illustrates a novel approach to opinion representation. Unlike the real values used in the continuous bounded confidence models to represent an opinion along some axis, opinions in the Axelrod model are of an enumerated type.

Voting Models

In voting models, which have been studied extensively since the seminal works of Clifford and Sudbury (1973) and Holley and Liggett (1975), each agent must make a binary choice between two candidates. The agents in these models behave like lemmings, randomly choosing one of their von Neumann neighbors and adopting its opinion. In these early models, the agents were arranged in a regular lattice and researchers found that the model was guaranteed to eventually reach consensus in one or two dimensions.

Suchecki et al. (2005) investigated the behavior of the voter model on different types of complex networks, and Castellano et al. (2007) found that the voter model does not converge to consensus on small-world networks.

Invasion Process What if, instead of meekly adopting the opinion of a neighbor, agents in the voter model pushed their opinion on a neighbor? This variation, called the "invasion process," was found by Sood et al. (2008) to be identical to the original form on regular latices, but different on irregular networks.

Drawbacks of Existing Models

Many of the existing models draw inspiration from social psychology but the underlying mathematical structures come primarily from physics. Indeed, these models are often grouped under the rubric "sociophysics." While the magnetic spin of atoms has been elegantly described by the Ising model, it's usefulness as a metaphor for human opinion is questionable.

1.2 Research Questions

This effort has been motivated by research questions pertaining to both domains and to the field of opinion dynamics more broadly.

STEM Interest

- **RQ1** What are the factors that cause interest to increase or decrease among the adolescents in a single urban community?
- **RQ2** Is the model described in this paper, which is based on current theories of interest development, sufficient to explain the trends observed in the data?
- **RQ3** What interventions might help foster higher STEM interest?

Radicalization

- **RQ4** How do psychological theories (which typically hypothesize intra-individual processes only) play out on the group level? Specifically, how does SQ theory work when implemented with an agent-based model of a simulated community?
- **RQ5** Under what conditions does it create group-level radicalization?
- **RQ6** Do different interaction topologies affect the amount of radicalization?
- **RQ7** What interventions does this ABM suggest to reduce or prevent radicalization?

General

- **RQ8** Is there a common framework that can be applied to opinion dynamics models in diverse domains?

1.3 Using ABMs to Find Answers

I answered these research questions by designing, building, and analyzing the STEM Interest model presented in chapter 3 and the Radicalization model presented in chapter 4. Both ABMs are based on the general model architecture described in chapter 2 and have rules based on the leading theories of subject matter experts in their respective fields. After verification and validation to establish model credibility, I ran parameter sweeps and experiments to assess the effect of different parameters over many simulation runs.

The models were both implemented in Java using the MASON simulation toolkit (Luke et al., 2005; Luke, 2015). MASON has the virtues of being flexible and fast, which allows for the systematic testing of different rulesets. Many of the existing models in opinion dynamics were implemented in custom native code Alizadeh and Cioffi-Revilla (2014); Alizadeh et al. (2014); Alizadeh and Cioffi-Revilla (2015); Alizadeh et al. (2015, 2016). Using a seasoned platform such as MASON provides well-tested building blocks that are known to work properly. This is a significant advantage in terms of model verification, and therefore, superior model credibility.

1.4 Main Findings

STEM Interest

RQ1 Interest in STEM is driven by peer and parental participation, passionate and knowledgeable adult leaders, and a high degree of choice in STEM-related activities. Interest declines when these factors are absent.

RQ2 The STEM Interest model described in chapter 3 accurately tracks the trends seen in the longitudinal data. The model is initialized with the 6th grade survey data, calibrated to match the 7th grade data, and successfully validated by comparing to the 8th grade data.

RQ3 Experiments with the STEM Interest model suggest two promising interventions. The first is brokering, i.e., when knowledgeable adults direct youth to information resources that suit their interests (see section 3.3.3). The second is increasing friend co-participation in group activities (see section 3.3.3).

Radicalization

RQ4 Supporting the motivational aspect of the SQ theory, there is empirical evidence from psychological studies showing that individual grievances lead to individual radicalization (Kruglanski et al., 2014), consistent with the multi-path theory (Cioffi-Revilla, 2010). The ABM in chapter 4 shows that individual grievances can also lead to (i.e., is a sufficient causal condition for) group radicalization providing initial support for the social aspect of SQ theory.

RQ5 The Radicalization model creates group-level grievance throughout the parameter space when traumatic events are frequent. Under the assumption that a community can be fully radicalized with an average of at least one event hitting each agent per step, the model is well-behaved for certain combinations of parameters shown in Table 4.2.

RQ6 Different interaction topologies do affect the amount of radicalization. In the absence of events, but with varying *latitude of non-commitment*, full-mixing produces more extremism due to increased pathways to the edge. With the model calibrated to data from Tetouan, Morocco, full-mixing produces fewer extremists than 8-Set and 4-Set when events are rare, but more extremists when events are frequent.

RQ7 The answer to RQ6 suggests that exposure to diverse opinions helps keep extremism low under normal conditions, but when the amount of extremism goes above a critical point, it may be beneficial to restrict communication networks.

General

RQ8 The general model described in chapter 2 has been successfully applied to opinion dynamics models in two significantly different domains (section 6.1 explores the extent of those differences).

The more complex nature of the answers to RQs 4–7 demonstrates the relatively greater complexity of radicalization as a social phenomenon compared to interest in STEM. However, the general model is shown to provide a viable scientific framework for both phenomena, so there is hope that it holds greater universal validity.

Chapter 2: General-Model

This chapter describes the *general model*, a framework designed to capture the common structure of all opinion dynamics models. section 2.1 describes the general model and its components, and section 2.2 shows how the general model has been applied to two significantly different domains.

2.1 Basic Framework

The basic framework of the general model is as follows. Agents have opinions which are affected by interactions with peers based on a social psychology paradigm and/or domain-specific knowledge. The form of the opinion and the response to the interaction will vary depending on the application. The components of the general model are shown as a UML diagram in Figure 2.1.

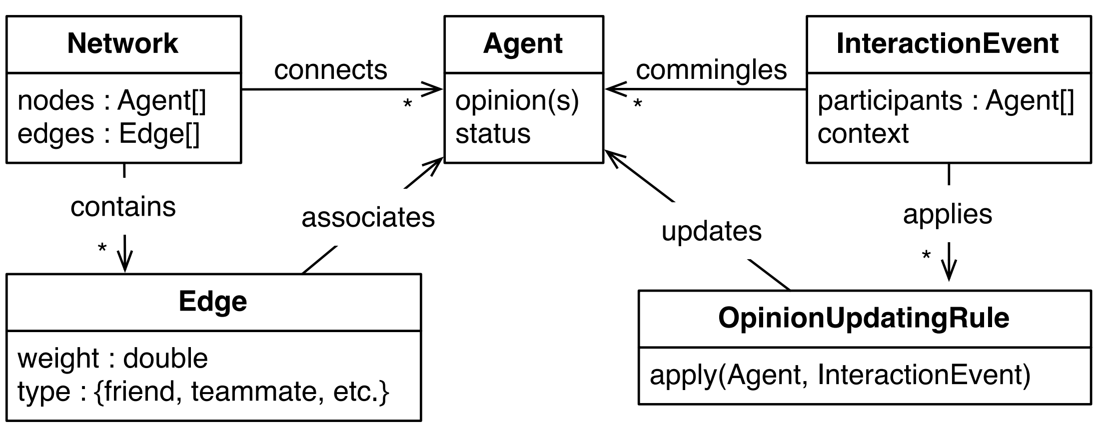


Figure 2.1: UML class diagram showing the components of the general model and their interactions.

The `Agent` class represents actors (e.g., citizens, students, etc.) in the model. Each agent has a set of `opinions`, which will vary throughout the simulation. The representation of an agent's opinions depends on the particular domain.

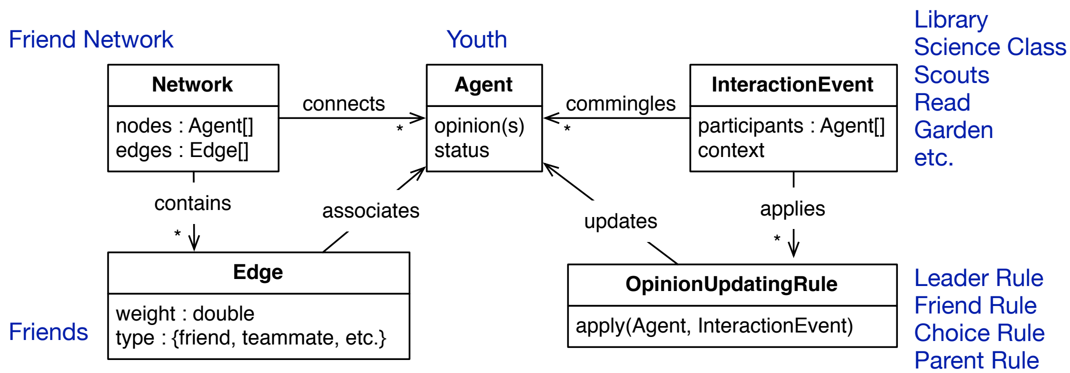
It has long been theorized that *power* is an important aspect of social influence (Emerson, 1962; French et al., 1959; Goldhamer and Shils, 1939). To that end, agent's may have a `status` attribute which may represent social standing, wealth, or reputation (Conte and Paolucci, 2002).

Social connections among agents will be stored in the `Network`, represented via `Edge`s. Each edge connects two agents with a `weight` that represents the strength of the relationship. This allows close friends to have more influence than acquaintances. Where necessary, the edge can include a `type` attribute describing the nature of the association (e.g., friend, teacher, family member, neighbor, etc.).

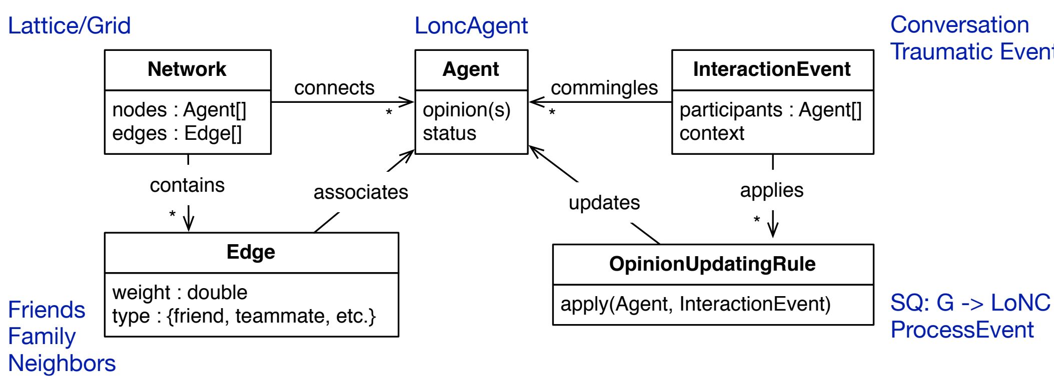
The opinions of agents will be changed during an `InteractionEvent` such as a traumatic event, after-school activity, etc. When one of these events occur, a list of `OpinionUpdatingRule`s will be applied that may change agent opinions in some way. For example, each agent's opinion may move toward the average opinion of the `participants` of the event à la Hegselmann and Krause (2002), or perhaps an agent becomes more interested in a topic because their friends are among the participants, or an agent is drawn toward extremism as a result of traumatic grievance.

2.2 Applications

The general model must be applied to two significantly different domains. By doing so, the general model's dynamics can be evaluated broadly and will be less susceptible to artifacts introduced by simplifying assumptions. The first application is a model of adolescents and their attitudes toward the fields of science, technology, engineering, and mathematics (STEM). The second application is a model of the radicalization of a community as they undergo traumatic events. Even though there is limited overlap between the domains and their particular implementations, they are both described elegantly using the general model framework, as shown in Figure 2.2.



(a) STEM Interest Model



(b) Radicalization Model

Figure 2.2: The STEM Interest and Radicalization models defined in terms of the general model.

STEM Interest

The primary agents in the STEM Interest model are the `Students` whose interest levels are being modeled. There are also `Adult` agents whose participation and attributes affect the `Students` but are not themselves the topic of investigation. `Students`' opinions are interest levels in four different topics, Earth/space science, Life science, Technology/engineering, and Mathematics. These interest levels are modified when the student participates in an `Activity`, which is the only `InteractionEvent`.

There are four `OpinionUpdatingRules`, the parent rule, choice rule, friend rule, and leader rule. Each of these is applied during when a student participates in any of the 21 different events.

Students are connected to their friends in the friend network, which affects participation and interest development. The `Adult` agents, who represent parents, teachers, and other leaders in the community, are not part of the friend network. Instead, they act as leaders during certain activities, and their attributes affect the outcome according to the parent rule and leader rule described in section 3.2.2.

Radicalization

Each agent in the Radicalization model has a political opinion ranging from -1 to 1, and a *latitude of non-commitment* (LoNC) value ranging from 0 to 2. Opinions are updated during interaction events (e.g., conversations) between pairs of agents. Interactions occur only between two agents who are connected to each other in the network, either a lattice or full-mixing topology.

The `OpinionUpdatingRule` of the interaction considers the distance between their opinions, and their respective LoNCs.

When a traumatic event impacts an agent, the agent's grievance goes up and its LoNC goes down, as described in the significance quest theory.

Part II

Domain Models

Chapter 3: STEM Interest: An Agent-Based Model of Interest Development Among Adolescents

Keywords: agent-based modeling, STEM, interest development, evolutionary computation, opinion dynamics

Acknowledgments: JFH and CCR are funded by the Center for Social Complexity at George Mason University. Major funding for the Synergies project was provided by the Noyce Foundation. These experiments were run on ARGO, a research computing cluster provided by the Office of Research Computing at George Mason University, VA. (URL:http://orc.gmu.edu)

This paper details a project using agent-based modeling to analyze the factors driving down interest in science, technology, engineering, and mathematics (STEM) during early adolescence and explores different strategies for improving outcomes. Several factors have hindered this type of research in the past, predominantly, the complexity of the mechanisms driving changes in youth interest in STEM, limited ability to study these mechanisms directly, and the difficulty of collecting data. We address these challenges by utilizing rich, longitudinal data collected over several years and developing a ruleset in conjunction with a panel of science education experts. We show that the hypothesized ruleset is sufficient to explain the trends observed in the data. We further show that interest may be retained/increased through knowledge brokering and increasing friend co-participation in STEM-related activities.

The paper is organized as follows. The introduction describes the motivation, research questions, and prior work. The method section describes the Synergies project generally, the data collection process, the design of the agent-based model, and the calibration process. The results section describes the calibration results, model validation, and a study of different intervention approaches. The discussion section describes the implications of our results, strengths and weaknesses of our model, and potential avenues for future work.

3.1 Introduction

American 4th graders, on average, have a competence and interest in science that ranks among the highest in the world. Likewise, American adults rank among the world's best in these same measures.[1] However, between 5th and 8th grades, American students undergo a precipitous drop in both interest and competence in all of the STEM fields (science, technology, engineering, and mathematics) (Falk and Dierking, 2010).

The Synergies project, headed by education researchers at Oregon State University and the University of Colorado at Boulder, is a hybrid research/intervention project designed to study the causes of this drop in interest in STEM, then design and deploy intervention strategies to help bolster interest.

The team selected the Parkrose neighborhood of Portland, Oregon as their area of interest. Parkrose is a place with many challenges. Bordered on the west by I-205, the south by I-84, and the north by the Columbia river, the neighborhood is largely isolated from the rest of the city. Portland's extensive public transit system stops at the edge; residents wishing to take transit downtown must first get themselves out of Parkrose before they can even catch a bus. Portland International Airport, only three miles away, brings constant noise pollution and drives down housing prices. Sixty percent of Parkrose residents are renters, many of whom receive government assistance. Comparing Parkrose to Portland overall, average home prices are $\approx$30% lower and the proportion of college graduates is $\approx$50% lower. Many of the residents are recent immigrants from eastern Europe who are not well engaged in the larger community. All four of the elementary schools in the Parkrose school district were rated under-performing in the 2011 No Child Left Behind evaluations.

[1] Whether this reflects well on Americans or poorly on the rest of the world is open to debate.

3.1.1 Motivation

The careers of the 21st century will increasingly be related to the STEM fields (science, technology, engineering, and math)(National Science Board, 2015). However, many children who might someday pursue a career in STEM lose interest during early adolescence. This trend, in which STEM interest among American adolescents declines dramatically between the ages of 10 and 14, is widespread and well-observed (George, 2006; Osborne et al., 2003; Simpson and Steve Oliver, 1990; Talton and Simpson, 1985; George, 2000).

The goal of this project is to use agent-based modeling (ABM) to analyze the factors driving this decline and evaluate potential remedies. We do this by building an ABM based on existing theories of interest development. The model utilizes rich longitudinal data collected over four years as part of the Synergies project (Falk et al., 2015a,b,c; Dierking et al., 2015; Penuel et al., in review).

3.1.2 Research questions

- What are the factors that cause interest to increase or decrease among the adolescents in a single urban community?
- Is the model described in this paper, which is based on current theories of interest development, sufficient to explain the trends observed in the data?
- What interventions might help foster higher STEM interest?

3.1.3 Background and Prior Work

The ABM presented in this paper models interest development as a function of STEM-related activities and how their circumstances shape the outcome. We draw on research showing that activities are more likely to foster interest when they are encouraged by parents (George and Kaplan, 1998), when they are done with friends (Podkul and Sauerteig, 2015), when they present a high degree of choice (Flowerday and Schraw, 2003), and when they are led by adults with expertise in the subject matter (Ebenezer and Zoller, 1993).

At a fundamental level, our ABM is based on the four-phase model of interest development of Hidi and Renninger (2006) which posits different levels, or phases, of interest. From low to high, these are: 1) *triggered situational interest*; 2) *maintained situational interest*; 3) *emerging individual interest*; and 4) *well-developed individual interest.* For our purposes, the primary distinction is between *situational* interest (phases 1–2), which relies on a stimulating environment or social cues from others, and *individual* interest (phases 3–4) which is sustained and leads a person to seek out new opportunities to learn. The ABM in this paper distinguishes between situation and individual interest using a threshold.

There has been surprisingly little work done using computational modeling to study interest in the STEM fields. Allen and Davis (2010) created a simple model based on social impact theory (Latane, 1981) to study the effect of peer conformity pressure on the yield of STEM majors. They found that placing talented STEM educators in classes for 9th and 10th grade generates 5.5% more yield than if they teach 11th and 12th grade classes. Sanchez et al. (2009) used a system dynamics model of STEM interest focused on teacher competence and turnover. Their work suggests that denying tenure to educators without STEM capabilities would, in time, increase the number of high school graduates choosing to purse a degree in STEM.

To date, there are no existing ABMs that model the role of STEM-related activities and their effect on interest. Nor are there ABMs studying the transition from situational to individual interest and vice versa. Nor are there ABMs that model a cohort of individuals over multiple years using longitudinal data, largely because data of this kind did not exist prior to this project. In short, this ABM is groundbreaking in many ways. Therefore, the MASON-based model demonstrated here provides a quantum leap in terms of theoretical, empirical, and overall analytical progress in terms of developing a deeper and policy-relevant understanding of the STEM crisis and ways to improve the situation.

3.2 Method

3.2.1 Project Overview

The Synergies project (Falk et al., 2015b, Falk et al. (2015c), Falk et al. (2015a), Dierking et al. (2015)) is a multi-year effort to investigate the processes that cause a decline in interest in the STEM fields among American adolescents. The project involves annual surveys of a cohort of students spanning an entire school district, beginning in fifth grade and ending in eighth.

Survey Description

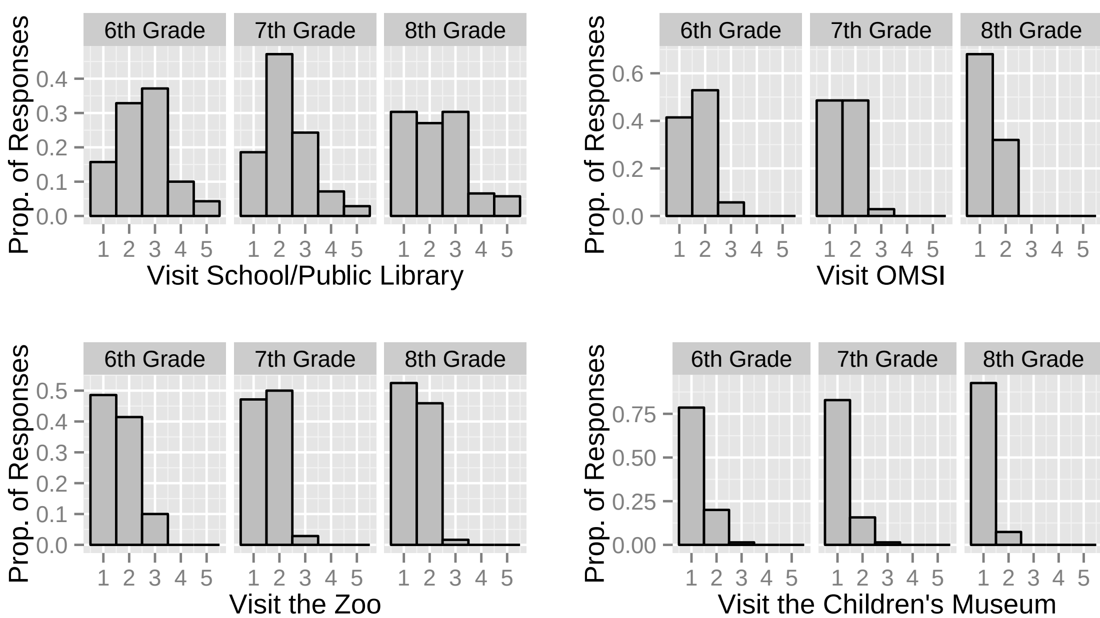


Figure 3.1: Proportion of responses to the first four questions about activity participation. (1="Hardly Ever or Never" to 5="Almost Every Day") The full set can be found in Appendix A

The survey includes several sections. The first section asks how often the respondent participates in a list of STEM-related activities like visiting the library, hiking, doing science kits, watching TV shows about STEM, etc. (see Appendix A for the full list and summary charts). Answers were given on a five-point Likert scale ranging from "Almost every day" to "Hardly ever or never." Figure 3.1 shows a summary of the responses to the first four activity questions. The full set can be found in Appendix A.

The survey also asked a series of questions to gauge the respondent's interest in a variety of STEM-related topics, e.g., how stars and planets are formed, what causes weather, how diseases work, etc. (see Appendix A for the full list). Answers to these questions were also given on a five-point Likert scale, this time ranging from "Like a lot" to "Dislike a lot." Falk et al. (2015d) describes the survey in detail.

3.2.2 Model overview

The model was implemented in Java using the MASON simulation toolkit (Luke et al., 2005; Luke, 2015). Each step represents one day, and on each day, each youth participates in 0–3 activities corresponding to those listed in the survey (see Appendix A), plus a `class` activity that takes place at school.

The model begins on September 4th, the first day of the 2012 school year. It's unusual for an agent-based model to be tied to a specific date like this, but it's useful in this case because some activities are only available during the school week, or the weekend, or only in summer, etc. Stepping through the days of a calendar provides a clear way to delineate all of this.

From Surveys to Model Input

The data provided by the surveys is richly detailed, but not immediately suitable for use in an agent-based model. For example, there are 16 items related to the respondent's interest in various STEM-related topics. Keeping these as separate interest levels would be unwieldy and would obscure correlations between the topics. To simplify, we used Principal Component Analysis (PCA) to reduce the 16 responses down to four topics. The topics, and the items composing them, are as follows:

Earth/space science (Cronbach's alpha=0.75):

- What it is like on other planets and exploring space
- How stars and planets form
- Why clouds, rain, and weather happen
- How earthquakes, volcanoes, and hurricanes happen

Life science (Cronbach's alpha=0.71):

- What to eat and how to exercise to keep healthy and fit
- How traits are passed from parents to children
- How the human body works

Technology/engineering (Cronbach's alpha=0.78):

- How buildings and bridges are made
- How computers, CDs, and cell phones work
- How to use and make maps
- How to design new games or toys
- How gas and diesel engines work

Mathematics (Cronbach's alpha=0.74):

- How to do Sudoku or other math problems
- How to measure the size or area of things
- How to solve puzzles
- How to make different shapes and patterns out of things

The resulting interest levels are written to a comma-separated values (CSV) file, with one row for each youth. However, not all respondents were included in the CSV file. As is typical when using survey data, invalid and incomplete responses have to be addressed during the data-cleaning process. In our case, we wanted to be able to track individuals over time, which meant we could only include respondents who completed both the 6th and 7th grade surveys. Table 3.1 shows the number of surveys collected each year and the portion of those with valid answers to all the questions used in the model. The last two rows show, respectively, the number of respondents who completed both the 6th and 7th grade surveys, and those who completed the 6th, 7th, and 8th grade surveys.

Although surveys were administered beginning in 5th grade, there were significant changes to the survey between the 5th and 6th grade samples. These changes were necessary improvements, but they made it infeasible to directly compare the data from 5th grade to those collected during 6th, 7th, and 8th grades. We therefore chose to omit the 5th grade data from this study.

Table 3.1: Number of surveys collected each year and the portion of those with valid answers to all questions used in this model.

Grade	Valid Surveys	Total Surveys
5th	129	174
6th	77	145
7th	97	162
8th	122	154
6th $\cap$ 7th	70	114
6th $\cap$ 7th $\cap$ 8th	52	90

Participation Rates

The survey asks how often the respondent participates in a list of activities and the responses are selected on a five-point Likert scale with the following options:

1. Hardly ever or never
2. A few times a year

3. 1-2 times a month
4. 1-2 times a week
5. Almost every day

These responses are converted to a `participationRate` reflecting the probability of participating in the activity at any given opportunity. This calculation is less trivial than it first seems. `Scouts/4-H`, for example, is only offered once per week, so an answer of "1-2 times a week" translates to 100% attendance. The code to calculate this transformation is shown in Listing 3.1.

Listing 3.1: Code mapping the Likert scale response from the survey to the participation rate.

```
double mapLikertToParticpationRate(int response) {
   int opportunities = 0;
   if (onSchoolDay)
      opportunities += 199; // a year minus weekends and summer
   if (onWeekend)
      opportunities += 104; // a year's worth of weekends
   if (onSummer)
      opportunities += 62;  // summer break minus summer weekends

   int timesDone = 0;
   switch (response) {
   case 1: timesDone = 0; break;        // Never
   case 2: timesDone = 3; break;        // few times per year
   case 3: timesDone = 18; break;       // 1-2 times per month
   case 4: timesDone = 78; break;       // 1-2 per week
   case 5: timesDone = 200; break;      // almost every day
   }

   return Math.min(1, timesDone / (double)opportunities);
}
```

Note that we ignore holidays and other breaks, so the number of school days in the model (199) is higher than the standard in real life (175-180).

Activity Types

In addition to the file containing initial youth data, we have a second file defining activity types. While the youth data file comes directly from the surveys (after some PCA), the activity types were defined by a panel of science education experts. The contents of the activityTypes.csv file are described in Table 3.2.

Table 3.2: Description of the columns of the CSV file defining activity types.

Column	Description
name	Name of the activity, shortened from survey
Earth/space science	Amount of earth/space science content
Life science	Amount of life science content
Tech./engineering	Amount of technology/engineering content
Mathematics	Amount of mathematics content
numLeaders	Number of unrelated adult leading this activity
numParents	Number of parents leading this activity
maxParticipants	Maximum number of participants
daysBetween	How often can a youth do this activity? 1 = every day, etc.
isOrganized	This is an organized activity (as opposed to ad hoc)
numRepeats	How many times is this done per session (organized only)
priority	Scheduling priority, 0 is highest (organized only)
onSchoolDay	This activity can occur on a school day
onWeekend	This activity can occur during the weekend
onSummer	This activity can occur during the summer
withFriendsOnly	Not doing this activity with strangers (ad hoc only)
degreeOfChoice	Participants choice in doing activity (high/moderate/low)

Some of these items (e.g., numRepeats, priority, onSchoolDay, onWeekend, and onSummer) are important for the scheduling mechanics of the model, but have limited relevance to theories of interest development. Others have direct implications.

The qualities of the adult leaders of an activity impact the youth's experience as described in the the leader rule. As such, `numLeaders` controls how many times the leader rule is invoked for an activity.

Parental participation affects the outcome of an activity according to the parent rule, and `numParents` determines whether this activity is generally done with a parent. Parental encouragement also affects the outcome of the parent rule, but encouragement varies from youth to youth according to their survey responses.

Organized activities have been found to promote experiences relating to "initiative, identity exploration and reflection, emotional learning, developing teamwork skills, and forming ties with community members." (Hansen et al., 2003). Organized activities also provide a regular opportunity to participate in activities with friends, which has a positive effect on interest development as described in the friend rule.

The `withFriendsOnly` flag, which applies only to ad hoc activities, determines whether a youth would only do that activity with friends (or alone). For example, a youth is only going to watch a STEM-related TV show with friends (i.e., `withFriendsOnly`=True), but the other youth at the zoo could be anyone. Activities that are done `withFriendsOnly` will receive more of a benefit from the friend rule.

As described in the choice rule, the `degreeOfChoice` affects whether the activity has a positive or negative effect on interest.

Activity Scheduler

Each step in the model represents one day, during which, each youth participates in some number of activities (not to exceed `maxActivitiesPerDay`). There are two categories of activities: organized and ad hoc. Organized activities occur on a regular basis with the same group of participants. These include `class`, `scouts/4-H`, `after-school-program`, `team-sport`, and `summer-camp`. The rest of the activities are ad hoc and occur stochastically when time permits.

Organized vs ad hoc activities

We distinguish between organized activities, which occur at a regular interval and include (some subset of) the same group of participants, and ad hoc activities, which occur sporadically whenever time permits. Research has shown that organized activities promote experiences relating to "initiative, identity exploration and reflection, emotional learning, developing teamwork skills, and forming ties with community members." (Hansen et al., 2003).

Organized activity groups are formed on the first day of each school year. Some organized activities have a lot of participants who have to be divided in multiple groups. Different approaches to group formation are studied in the results section. Anyone with a participation rate greater than zero will be assigned to a group for that activity. When the activity occurs, they may or may not join based on their participation rate.

The differences between organized and ad hoc activities are as follows:

Organized:

- Scheduled before ad hoc
- Scheduled by priority
- Occurs at regular interval
- Repeated a fixed number of times
- Groups formed on first day of the school year
- Different grouping methods explored in results section
- Youths have some probability of participating each opportunity
- Youths who report that they never do the activity aren't assigned to groups

Ad hoc:

- Can occur whenever schedule permits
- Limited by `daysBetween`
- Order randomized each day before scheduling to avoid bias (no priority)
- Youths have some probability of participating each day

- May join others for activity (depending on type)

Daily activities

Each day, a set of activities is selected for each youth. Since there's a limit on the number of activities per day (3 by default), it's necessary to consider them in order of priority. We wouldn't want, for example, a youth to fill their quota with `reading`, `gardening`, and `hiking`, and not have time to attend `class`. Organized activities are scheduled first, with `class` and `summer-camp` having the highest priority during the school year and summer, respectively, followed by a tie between `team-sport` and `after-school-program`, and lastly followed by `scouts/4-H`.

If there's still room in the youth's schedule, they may add some ad hoc activities, which are considered in random order. If they choose to do an activity that can includes friends, they look to see if any of their friends want to join them.

Rules

Interest levels are modified by a series of rules. When an activity occurs, each rule is evaluated for each participant. The logic of the rule determines whether a youth's interest level increases, decreases, or remains unchanged. When the interest level is changed, it is done according to Equations 3.1, and 3.2:

Increase Interest:

$$\text{interest}_{topic} = \text{interest}_{topic} + \delta * \text{relevance}_{topic} * \text{weight}_{rule} \tag{3.1}$$

Decrease interest:

$$\text{interest}_{topic} = \text{interest}_{topic} - \delta * \text{relevance}_{topic} * \text{weight}_{rule} \tag{3.2}$$

Where:

interest_{topic} is the youth's interest in a topic

δ is the system-wide interest change rate

relevance_{topic} is the relevance of the current activity's content to the topic

weight_{rule} is the weight of the current rule

Parent Rule

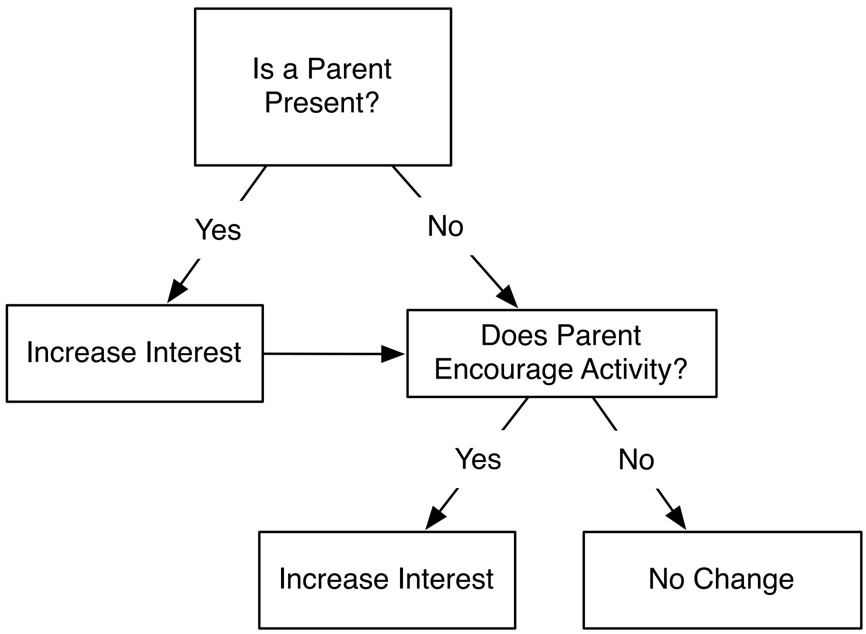


Figure 3.2: The Parent rule

(George and Kaplan, 1998) found that parental involvement is very important in fostering positive attitudes toward science. Accordingly, if a parent is present for an activity, or encourages that activity, interest increases.

Choice Rule

Flowerday and Schraw (2003) showed that having choice in an activity has a positive effect on attitude and effort. To capture this, each activity is rated as having a low, moderate, or high degree of choice. Interest decreases when the degree of choice is low, increases when it's high, and remains unchanged when it's moderate.

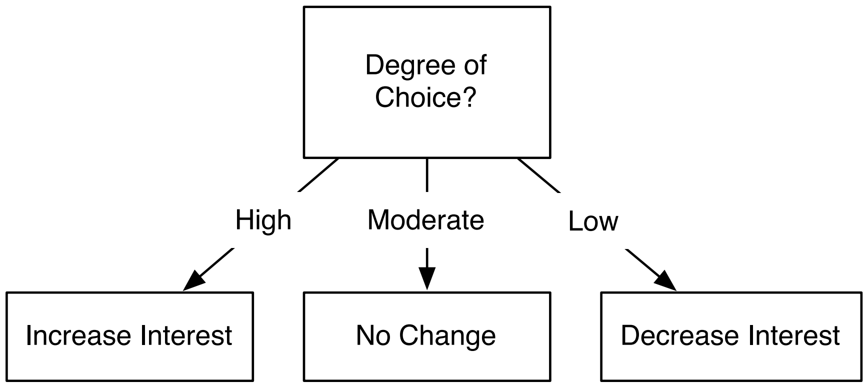


Figure 3.3: The Choice rule

Friend Rule

If an activity is done with a friend, interest increases. If there are no friends present and the youth is below the interest threshold, interest decreases. Otherwise, there's no change.

Leader Rule

The importance of expertise in activity leaders was demonstrated by Ebenezer and Zoller (1993), who showed that interest levels among 10th graders increase when they perceive their teachers as experts. Members of the study community suggested that passion is also an important factor, though it plays a different role.

Some activities are led by adult leaders and the expertise and passion of the leader affects the outcome. If the leader has passion and expertise, interest increases. If the leader has neither passion nor expertise, interest decreases. If the leader has passion but low expertise, interest

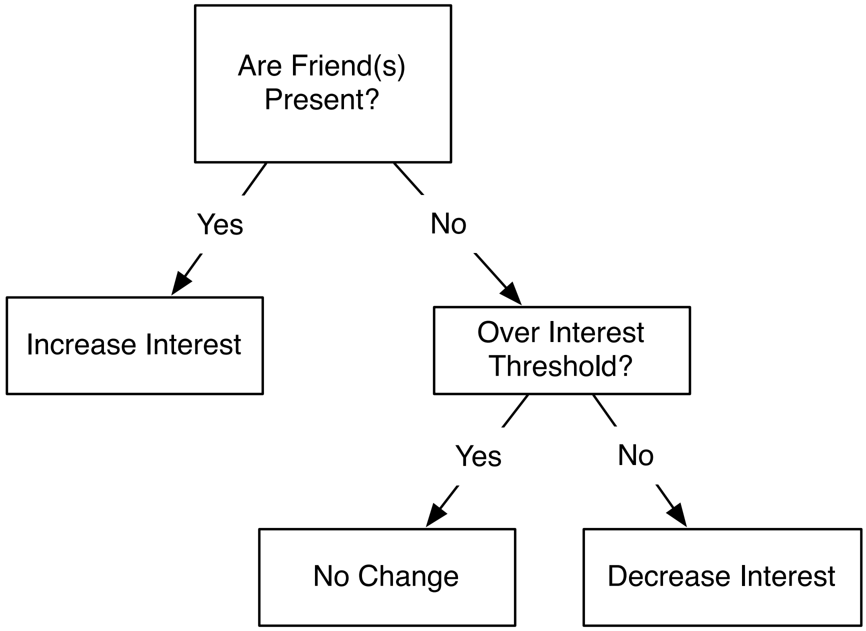


Figure 3.4: The Friend rule

increases for students who are below the interest threshold and decreases for those above the threshold. If the leader has expertise but low passion, interest increases for students who are above the threshold and decreases for those below it.

Friend Network

The youth are connected in a small-world network generated by the (Watts and Strogatz, 1998) algorithm. The algorithm works by first forming a ring network, then randomly rewiring links. We assume an average of 3 friends per student based on the research of Xie et al. (1999), and a rewire probability of 0.5. To make friendships between boys and girls relatively rare, they are formed into separate small world networks, with random rewiring between them. For the purposes of this study, boys and girls aren't friends (i.e., they exist within separate friend networks).

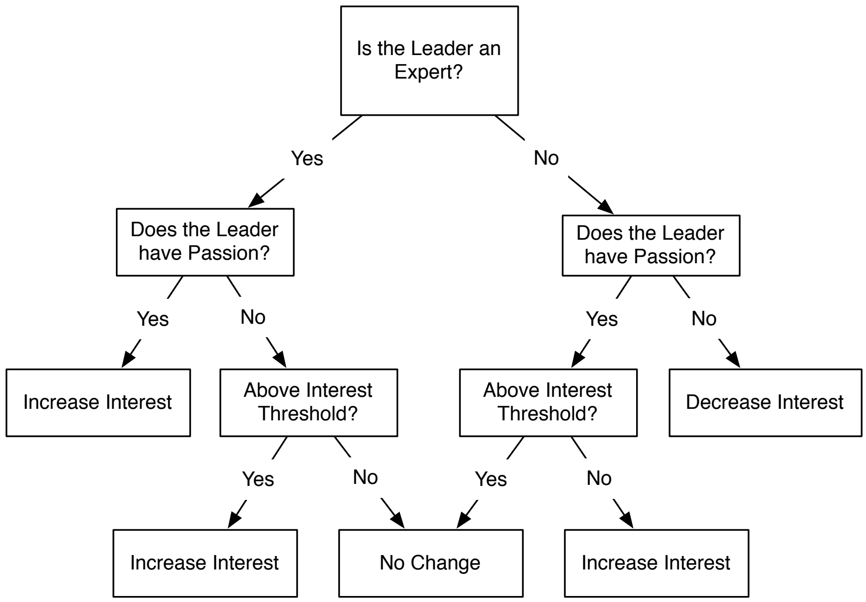


Figure 3.5: The Leader rule

Example

It is a Monday during the school year. Maria goes to school and participates in the `class` activity. She has a friend in her class, which increases her interest in the topic according to the `friend-rule`. However, she is dragged down by the low degree of choice in the topic, which reduces her interest via the `choice-rule`. Her teacher isn't very passionate about the topic, but is an expert. Since Maria already has an individual interest in the topic, she cares more about expertise than passion and her interest goes up due to the `leader-rule`. Her parents are very encouraging about her school studies, so her interest increases via the `parent-rule`.

After school she has soccer practice, which she attends 80% of the time. However, this is one of those rare occasions when she doesn't attend.

Instead, she decides to go home and do a `science-kit`. She asks around, finds that her friend Sarah is also planning to do a `science-kit`, and decides to join her. The `friend-rule` increases her interest. The `choice-rule` is neutral due to the activity's moderate degree of choice. Her parents don't particularly encourage her to do science kits, so the `parent-rule` has no effect, and the `leader-rule` doesn't apply.

She heads home from Sarah's house and gets on the computer to find out more about the experiment they just performed (i.e. `visit-website`). This is a solitary activity (e.g. `maxParticipants` = 1), so the `friend-rule` doesn't apply. She enjoys being able to choose what sites she visits, and the high degree of choice drives up her interest via the `choice-rule`. Her parents don't offer much encouragement when she's online, so the `parent-rule` is neutral. As with the previous activity, the `leader-rule` doesn't apply.

After doing these three activities, she has no time for anything else STEM-related (since `maxActivitiesPerDay` = 3).

3.2.3 Calibration

Even with a well-chosen set of rules, a model must be calibrated before it will produce output similar to real-world data. With a model as complex as this one, with dozens of parameters, that can be particularly challenging. One popular method of calibrating a model with a large parameter space is to use evolutionary computation (De Jong, 2006; Calvez and Hutzler, 2005; Stonedahl and Wilensky, 2010), a process analogous to Darwinian natural selection.

The algorithm creates a population of individuals who compete against one another and the fittest among them survive and reproduce. In this case, each individual is a set of model parameters, and its fitness is a measure of how accurately the model runs with those settings. Each generation the fittest individuals are identified, then duplicated, mutated, and crossed-over to produce offspring (i.e., new sets of parameters). In a surprisingly small number of generations, often only a couple dozen, the algorithm homes in on the optimal solution.

Fitness

Fitness is determined by comparing model output to survey data. Specifically, we load the 6th grade survey data into the model, run it for a year of simulated time, then compare the state of the model to the 7th grade survey data. We compare the distributions of interest levels for each topic along with the distributions of activity levels for each activity. Goodness-of-fit is calculated using the Kolmogorov-Smirnov (KS) statistic (Massey Jr, 1951), which has the benefit of being nonparametric and thus easy to combine. The KS statistics for the four topics are averaged, as are the KS statistics of the twenty topics, then those averages are averaged to calculate the fitness for that simulation run (see Equation 3.3).

$$f = \frac{1}{2}\left[\frac{\sum_{t\in\text{Topics}} \text{KS}(I_t^{\text{model}}, I_t^{\text{data}})}{\#\text{Topics}} + \frac{\sum_{a\in\text{Activities}} \text{KS}(P_a^{\text{model}}, P_a^{\text{data}})}{\#\text{Activities}}\right] \tag{3.3}$$

Where:

f is fitness

$\text{KS}(x, y)$ is the Kolmogorov-Smirnov statistic between distributions x and y

I_t^{model} is the distribution of interest in topic t in the model

I_t^{data} is the distribution of interest in topic t in the data

P_a^{model} is the distribution of participation in activity a in the model

P_a^{data} is the distribution of participation in activity a in the data

Agent-based models with stochastic elements typically have some variation from run to run. This potentially introduces noise into the fitness score of the individual (e.g., set of model parameters). To produce a more accurate fitness evaluation, the standard practice in evolutionary computation is to run the simulation multiple times and average the fitness scores. This ensures that an individual's fitness reflects its consistent performance rather than a fluke. In this project, each individual's fitness is the average of five runs.

EC Specifics

For calibration, we used the Evolutionary Computation Java (ECJ) toolkit (Luke et al., 2007; Luke, 2011). The population contains 100 individuals, each of which is a genome with 13 parameters (described in detail in the Results section). Each successive generation is selected via a two-player tournament, in which two random individuals are selected, evaluated, and compared. The individual with higher fitness survives and is paired with another tournament winner to reproduce. The two are then copied, crossed-over with two-point crossover, and mutated by adding Gaussian noise with a standard deviation of 0.05. This is repeated 50 times to produce the 100 offspring that compose the next generation. After 50 generations, ECJ terminates and outputs the fittest individual in the whole run.

Complex fitness landscapes often have multiple peaks and it's possible for the evolutionary algorithm to converge on one of the suboptimal peaks (De Jong, 2006). To increase the odds of finding the optimal peak, we performed 50 separate ECJ runs and used the fittest individual of all.

3.3 Results

3.3.1 Calibration

The model contains parameters controlling the interest threshold, the rates at which interest and participation levels change, the characteristics of adult leaders, and the weights for each of the interest development rules. Table 3.3 shows the calibrated values for each of these parameters and the constraints imposed on their range.

Interest Threshold

One of the key concepts in this model is the distinction between situational and individual interest (Hidi and Renninger, 2006). When a person's interest in a topic is relatively low,

Table 3.3: Model parameters calibrated via evolutionary computation to fit the longitudinal survey data.

Parameter	Calibrated Value	Constraints
InterestThreshold	0.94842540355904636	0.85 - 0.95
InterestThresholdSD	0.00946183819792444	0 - 0.5
InterestChangeRate	0.001	fixed
ParticipationChangeRate	0.0	fixed
ParticipationMultiplier	1.0	fixed
LeaderExpertise	0.18071591495957151	0 - 1
LeaderExpertiseSD	0.06630778916338272	0 - 1
LeaderPassion	0.11527919391589679	0 - 1
LeaderPassionSD	0.07878072569768096	0 - 1
FriendRuleWeight	0.20004421974661746	0 - 1
ChoiceRuleWeight	0.23159159447967126	0 - 1
ParentRuleWeight	0.32469239696194879	0 - 1
LeaderRuleWeight	0.47357174957063619	0 - 1

they are drawn in primary by situational factors like encouragement from parents or friends, the passion of adult leaders, etc. When their interest is high, they are individually driven to learn more about the topic and the social aspects of an activity become secondary to the informational content. The interest threshold parameter represents the theoretical boundary between situational and individual interest.

The survey data provides us with the interest levels of our population, but it doesn't tell us which youth are individually interested in which topics. In retrospect, this would have been a valuable addition to the survey. Without that data, we are left to calibrate the interest threshold with imperfect information. One useful metric we can inspect while adjusting the threshold is the percentage of youth who are above it in at least one topic. With the threshold set at 0.5, we find 95% of the youth interested in a STEM topic. We know from Maltese and Tai (2011) that 16% of students choose to pursue STEM majors in college, and while this isn't exactly the same metric, nor is it measuring youth of the same age, it's enough to suggest that our 95% figure is far too high. It lacks face validity. We're assuming the correct value to be somewhere in the neighborhood of 20%.

Constraining the range of the interest threshold to be between 0.85 and 0.95 produces runs in which roughly 20-35% of the youth are interested in a STEM topic. The optimized threshold (0.948) results in 24% of youth interest in STEM.

The standard deviation of the interest threshold was allowed to range from 0 to 0.5, but the optimized value was very small (0.009). This is consistent with the other variation parameters which all settled on small numbers. By driving the variation down, the evolutionary algorithm is reducing the heterogeneity among the agents. This suggests that it's easier to optimize this model with little or no variation among the agents.

Interest Change Rate

This parameter affects the amount interest changes whenever it goes up or down. Its effect is system wide. When a rule executes and increases (or decreases) interest, the weight of the rule is multiplied by this value to determine the size of the interest change, as shown in the section 3.2.2. As such, it is mathematically redundant and could be removed without a loss of functionality. However, keeping it and fixing it to 0.001 yields calibrated rule weights in the range of 0.2 to 0.5 rather than 0.0002 to 0.0005, which is arguably more convenient to work with. In either case, the parameter doesn't need to be calibrated, which is why it is fixed.

Participation Change Rate

This parameter controls the rate at which participation changes whenever it goes up or down. However, the survey data shows almost no change in the reported activity levels between 6th and 7th grades. The calibration runs consistently drove this value lower and lower since doing so resulted in a closer fit to the data, and thus a higher fitness score. Rather than artificially constraining the model to ensure some non-zero level of participation change, we decided to set this to zero and hold participation levels constant throughout the run. This forced the evolutionary algorithm to find improvements in fitness through the parameters related to interest development.

Participation Multiplier

Over-reporting of positive traits (and under-reporting of negative ones), called *social desirability bias* and is a widely-observed problem with surveys in general(Fisher, 1993). This parameter allows for a system-wide adjustment of activity levels to potentially correct for over-reporting in the survey data. For the purposes of calibration and the experiments in this paper, we're assuming that the participation levels reported in the surveys are accurate. Thus, this parameter is fixed at 1.

Leader Expertise and Passion

The leader rule dictates that the expertise and passion of the adult leader of an activity impact the experience of the participants. For youth with individual interest in a topic, expertise is more important, while youth with only situational interest are affected more by a leader's passion.

Leaders with passion above 0.5 are considered passionate, and those with expertise above 0.5 are considered expert. The expertise level for each leader is drawn randomly from a normal distribution with a mean equal to `LeaderExpertise` and a standard deviation equal to `LeaderExpertiseSD`. The passion levels are similarly determined by `LeaderPassion` and `LeaderPassionSD`.

The calibrated values found by the evolutionary algorithm are small (expertise ≈ 0.18 with SD ≈ 0.07, and passion ≈ 0.12 with SD ≈ 0.08). These numbers are small enough that fewer than one leader in a million will have passion or expertise over 0.5. Essentially all the leaders in this model are low passion and low expertise.

This shouldn't be taken to suggest that there are no knowledgeable or passionate adult leaders in the study community. Rather, it indicates that the evolutionary algorithm was able to fit the target data more closely when it drove these parameters all the way down. The leader rule is more complicated than any of the other rules. Its outcome depends on two characteristics of the leaders, and may differ depending on the interest level of the participant. Moreover, it handles each topic separately, so a youth's interest in biology could go up during the same activity that his/her interest in math goes down. By setting the passion and expertise parameters effectively to zero, all that complexity is bypassed and the effect of the rule is always a decrease in interest. This is one of the interesting aspects of using evolutionary computation to calibrate models. If there is a shortcut to high fitness, the algorithm will find it and exploit it.

Rule Weights

The rule weights affect how large of an impact each rule has when it increases or decreases interest. The leader rule has the largest weight at 0.47, and as we observed in the previous section, its impact is always negative. The next largest is the parent rule at 0.32, followed by the choice rule at 0.23, and the friend rule at 0.20. The relative strength of these weights can be seen in Figure 3.6 which shows the net effect of each rule on the population. Note that the leader rule has the largest impact.

3.3.2 Validation Against 8th Grade Surveys

The calibration process for this model involved reading the 6th grade survey data, running the model for a year of simulation time, and then comparing interest levels in the model to the 7th grade survey data (see the subsection 3.2.3 for details). The result is the set of calibrated model parameters described in subsection 3.3.1. When the model is run with these parameters, it produces a simulated population whose interest levels closely match those in the 7th grade survey data. By running the model for another year and comparing

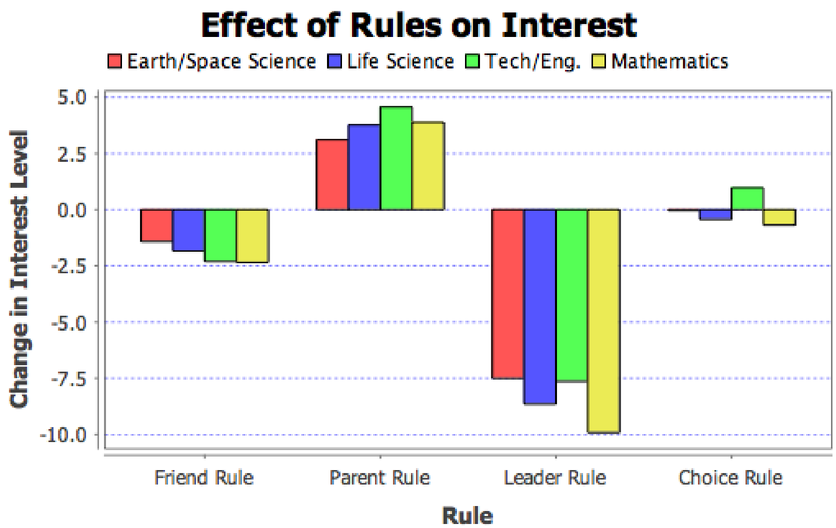


Figure 3.6: Net effect of rules on all youth's interest levels in the calibrated model. The friend rule and leader rule have a negative impact, while the effect of the parent rule is positive overall. The impact of the choice rule is mixed.

against the 8th grade survey data, we have a powerful validation test. If the fit is good, we'll have shown that the calibrated model has predictive capability. If the fit is not good, it suggests that we've overfit the data. The results are shown in Figures 3.7, 3.8, 3.9, 3.10.

Table 3.4 shows the results of the Kolmogorov-Smirnov test comparing the model output to the corresponding survey data. All the KS tests, with one exception, failed to reject the null hypothesis that the samples are drawn from the same population. The exception is life science in 8th grade ($p = 0.0008$), which saw an uptick in interest since the 7th grade survey (see Figure 3.8), whereas interest in the other topics declines between 7th and 8th grade.

3.3.3 Interventions

Now that we've established that our model does a reasonable job of capturing the underlying dynamics of interest development, we naturally want to use it to improve outcomes. Computational modeling provides an opportunity to experiment with different intervention approaches *in silico* without the cost or ethical complications of intervening *in vivo* (Epstein,

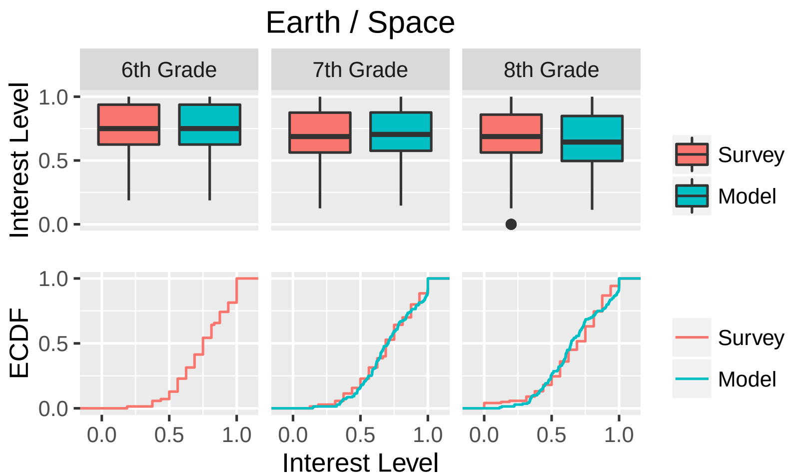


Figure 3.7: Comparison of interest in earth/space science in the survey data and model output.

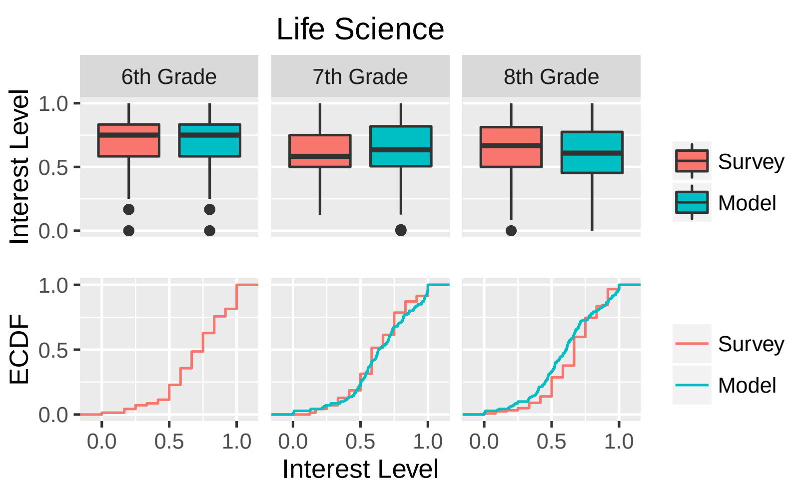


Figure 3.8: Comparison of interest in life science in the survey data and model output.

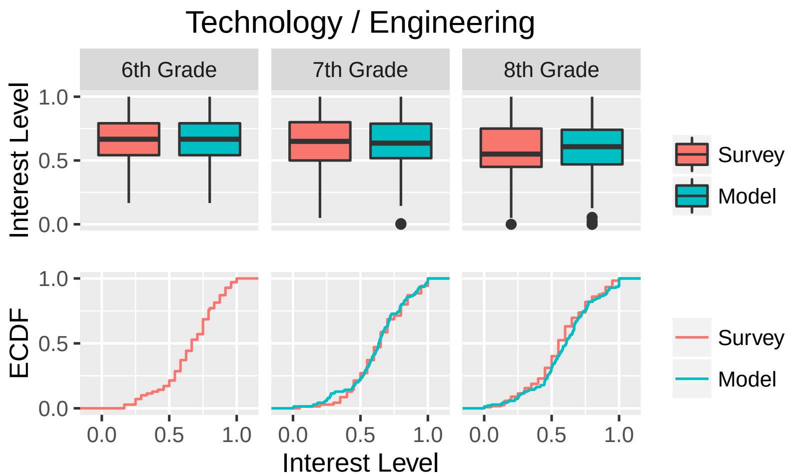


Figure 3.9: Comparison of interest in technology/engineering in the survey data and model output.

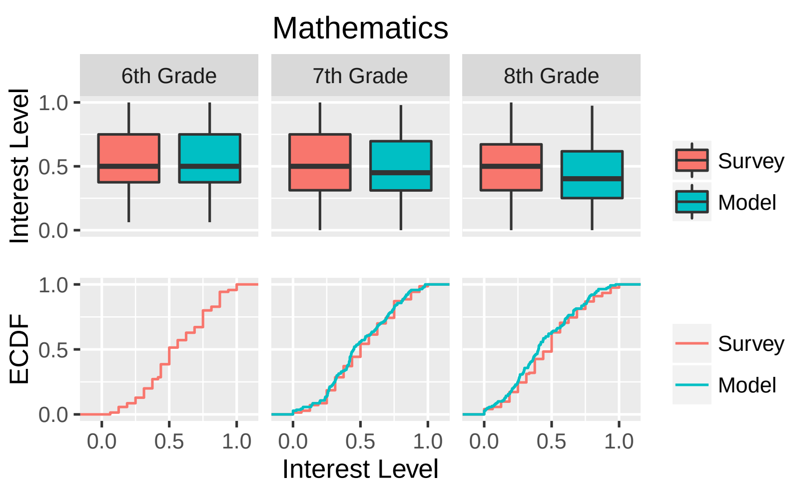


Figure 3.10: Comparison of interest in mathematics in the survey data and model output.

Table 3.4: Validation of model output against survey data. In every case except one (in bold), the Kolmogorov-Smirnov test failed to find a significant difference between model output and survey data.

	7th Grade		8th Grade	
Topic	KS statistic	p-value	KS statistic	p-value
Earth/Space Science	0.0786	0.7805	0.1622	0.0648
Life Science	0.1143	0.3200	**0.2444**	0.0008
Technology/Engineering	0.0857	0.6826	0.1454	0.1268
Mathematics	0.1143	0.3200	0.1521	0.0979

2008). This section presents several such approaches and evaluates them for their potential efficacy.

Brokering

Competent knowledge brokers can aid interest development by directing inquisitive young people toward resources that align with their interests (Barron et al., 2009; Ching et al., 2015). In the model, we capture this phenomenon with a simple mechanism. When a youth meets with their broker, the broker gives them a website to visit later that day. Specifically, a `visit-website` activity is scheduled that contains content focused on the youth's primary interest.

For example, a youth with an interest vector of $< 0.2, 0.1, 0.4, 0.6 >$ (i.e., primarily interested in mathematics) would be directed to a website with a topic vector of $< 0, 0, 0, 1 >$ (i.e., focused on math). Website activities have a high degree of choice, resulting in an increase of interest per the choice rule.

The brokering mechanism is controlled by two parameters. The first, `proportionWithBroker`, controls how many youth have brokers. The second parameter, `brokerProbability`, determines the probability that a youth (who has a broker) will talk to their broker on any given day. A `brokerProbability` of 0.2 indicates a 20% chance of brokering on any given day, i.e., once every five days on average.

Figure 3.11 shows that increased brokering improves average interest levels and the difference is statistically significant. However, the scale of the change is small—only increasing average interest from 0.618 to 0.624 for a value that ranges from 0 to 1, an increase of approximately one percent.

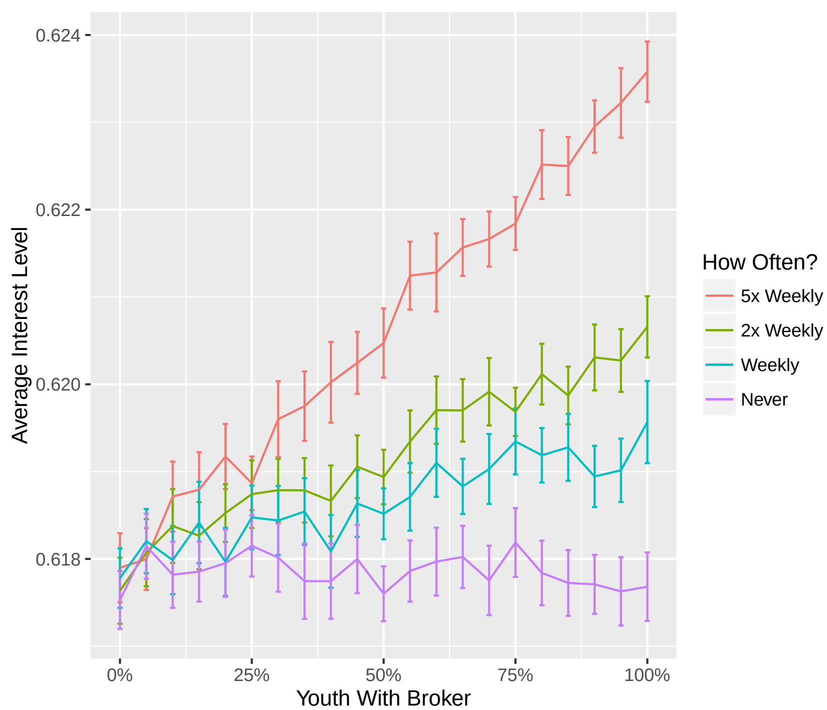


Figure 3.11: The effect of brokering on the average interest of the population. Confidence intervals are shown at 95%.

Although the change in average interest is small in magnitude, it is focused for each youth in the topic that interests them most. This results in a much more profound effect on the percentage of youth who have an individual interest in their favorite STEM topic. Figure 3.12 shows a 42% increase in the percentage of youth with an individual interest in STEM, from ˜24% to almost 34%.

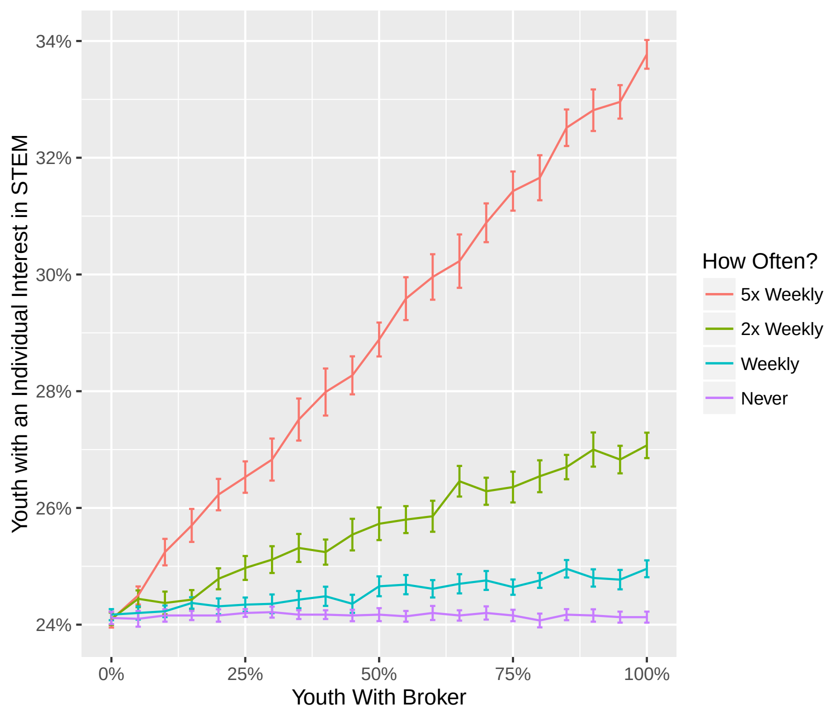


Figure 3.12: The effect of brokering on the percentage of youth who have an individual interest in a STEM topic. Confidence intervals are shown at 95%.

It's clear that increased brokering improves the outcome in our model, but how should the brokers allocate their efforts to make the greatest impact? Given practical time constraints, is it better to meet infrequently with every youth, or frequently with only a few? The answer depends on your objective.

For the sake of analysis, let *Brokering Resources* be the product of `proportionWithBroker` and `brokerProbability`. This represents the total amount of time and other resources spent on brokering. If it takes x resources to broker for one youth monthly, it takes $2x$ to broker for two of them, and $4x$ to do so twice monthly. Since `proportionWithBroker` and `brokerProbability` are both in the range 0-1, Brokering Resources is too.

Similarly, let an *allocation* be a unique combination of `proportionWithBroker` and `brokerProbability`. For example, Brokering Resources of 0.16 could be divided up in the allocations shown in Table 3.5.

Table 3.5: Example allocations of brokering resources.

proportionWithBroker	brokerProbability	Brokering Resources
0.2	0.8	0.16
0.4	0.4	0.16
0.8	0.2	0.16

We performed a full sweep of the brokering parameters, `proportionWithBroker` and `brokerProbability`, in increments of 0.05, running the model 50 times at each allocation. Each dot in Figure 3.13 represents the average of those 50 runs. The black circles highlight the Pareto-optimal allocations of resources. For a given quantity of resources, these are the best allocation of `proportionWithBroker` and `brokerProbability` for maximizing average interest level.

Figure 3.14 shows the parameter sweep in two-dimensions. The color gradient shows, as we saw in Figures 3.11, 3.13, that increasing brokering increases average interest. The

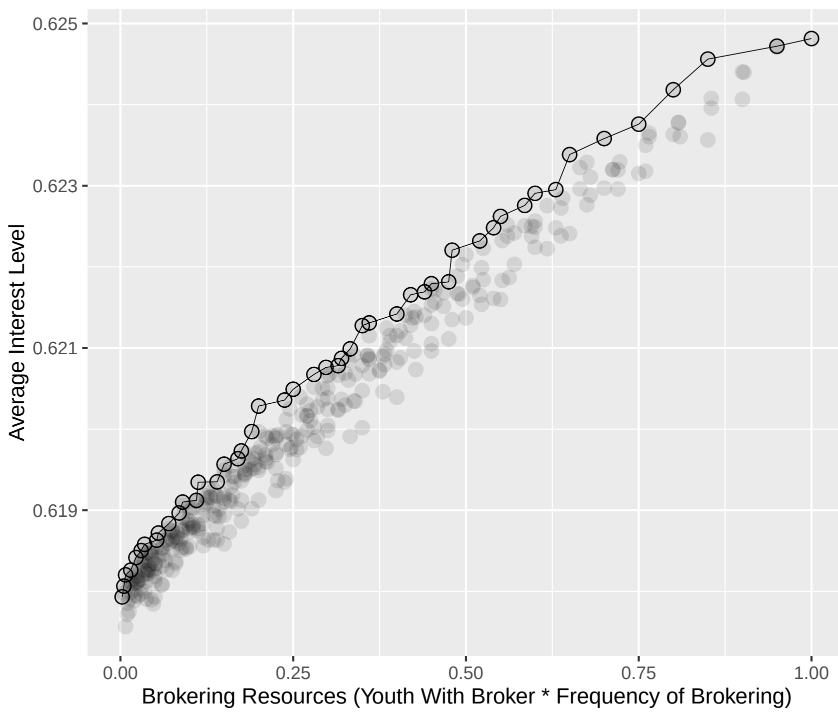


Figure 3.13: Brokering allocations, ordered by resources, vs. average interest. Each dot represents the average of 50 runs at one allocation set point. Pareto-optimal allocations are circled in black. Increasing brokering resources from 0 to 100% increases average interest from 0.618 to 0.625.

circles highlight the same Pareto-optimal allocations circled in Figure 3.13. The hyperbola shows where Brokering Resources equals 0.05. Above this threshold, all the Pareto-optimal allocations result in a statistically significant (with 99% confidence) increase in average interest level over the runs with no brokering.

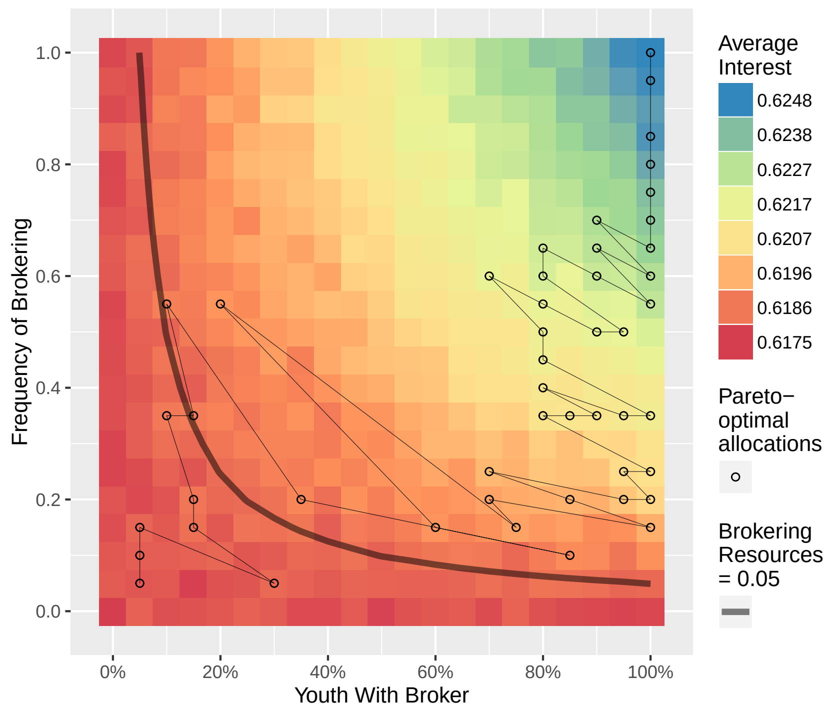


Figure 3.14: To increase average interest, spread brokering thin. Circles show Pareto-optimal allocations. Heavy line indicates where Brokering Resources = 0.05, above which all Pareto-optimal allocations are significantly better than zero brokering.

Although there's a fair amount of noise, especially in the lower-left half, a dominant pattern eventually emerges. The most efficient way to increase average interest through brokering is to broker for a large portion of the youth infrequently.

When the objective is to increase the number of youth who have an individual interest in STEM (which we can think of as likely STEM majors), the situation changes. Figure 3.15 shows all the allocations in the parameter sweep, ordered by Brokering Resources. The

Pareto-optimal allocations are circled in black.

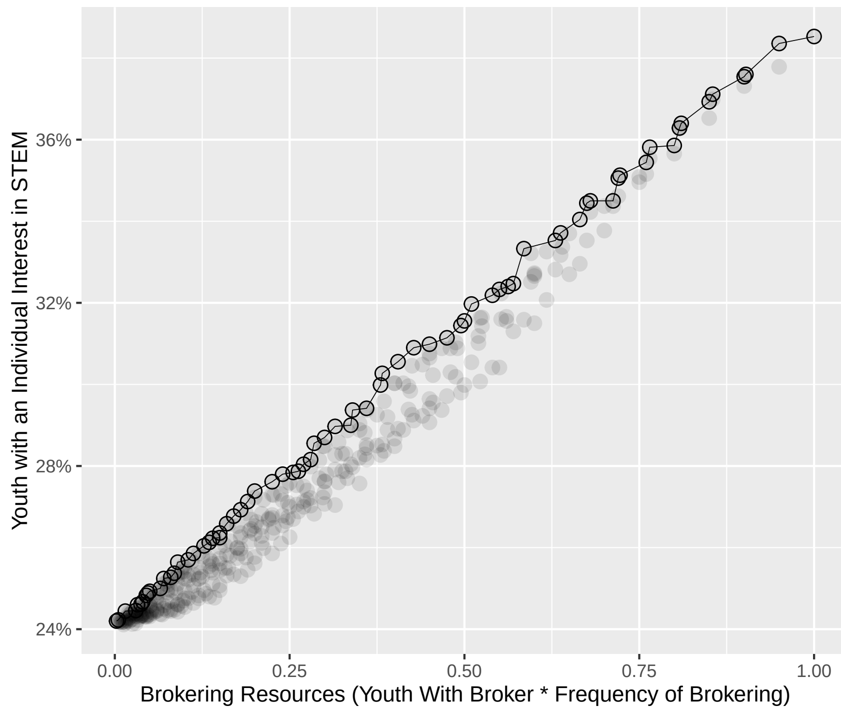


Figure 3.15: Brokering allocations, ordered by resources, vs. the percentage of youth with an individual interest in STEM. Each dot represents the average of 50 runs at one allocation set point. Pareto-optimal allocations are circled in black. Increasing brokering resources from 0 to 100% increases the percentage of youth with an individual interest in STEM from 24% to 34%.

Overlaying the Pareto-optimal allocations on the 2D tile plot in Figure 3.16, we see that the most efficient way to increase the percentage of youth with an individual interest in STEM is to choose a small group and broker for them frequently. This is somewhat analogous to gifted and talented programs, except the brokered youth in the model are chosen randomly.

Friend co-participation (grouping)

At the start of every year in the model, youth participating in an organized activity (e.g., `scouts/4-H`, `after-school-program`, `team-sport`, or `class`) are divided into groups. For

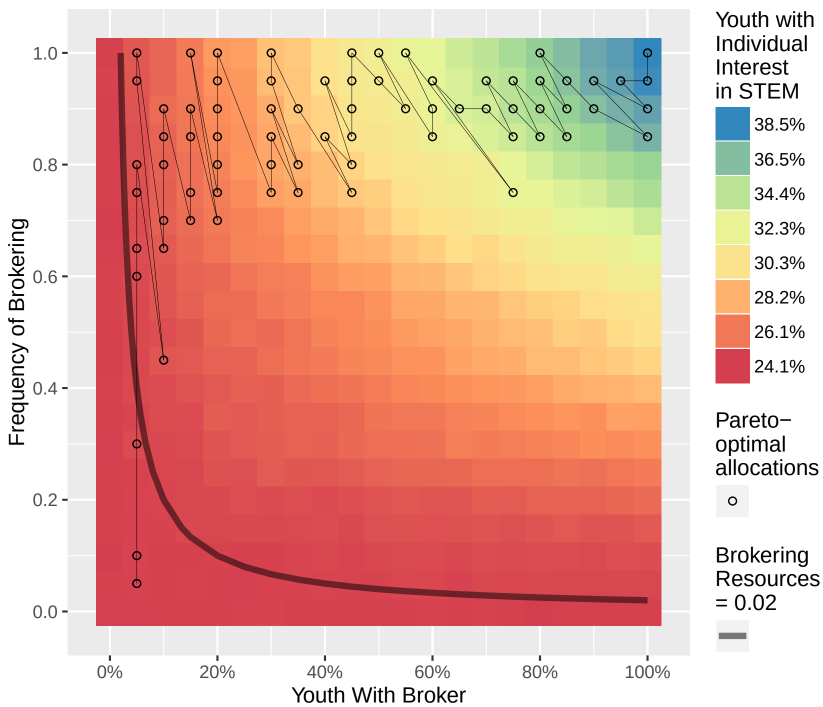


Figure 3.16: To get more STEM majors, focus brokering on a small group. Circles show Pareto-optimal allocations. The hyperbola indicates where Brokering Resources = 0.02, above which all Pareto-optimal allocations are significantly better than zero brokering.

example, the maximum size of an `after-school-program` group is 30 participants, so if more than 30 youth sign up, they are divided into separate groups. The model contains three different grouping methods that affect how likely youth are to be grouped together with their friends.

The first method, called `Random`, simply chooses a group at random for each participating youth. The second method is called `InOrder` and fills the first group until it is full, then the second, and so on. As shown in Table 3.6, this results in more friend co-participation than `Random` because the youth are initially ordered by school teacher and the friend network formation method makes friendships more likely between classmates. The third method, called `FriendChain`, adds one youth and all their friends (assuming they also participate in the activity in question), followed by their friends' friends, and so on. It does this in a depth-first traversal. The code for `FriendChain` is shown in Listing 3.2:

Listing 3.2: Code that groups youth according to the FriendChain method.

```
/**
 * Match participants to activities by adding a participant,
 * then his/her friends, then their friends, etc. Uses a
 * depth-first traversal.
 */
void groupByFriendChain(ArrayList<Student> participants,
    ArrayList<Activity> activities) {
    for (Activity a : activities) {
        while (!a.isFull() && !participants.isEmpty())
            addFriendChain(participants.get(0), a, participants);
        if (participants.isEmpty())
            break;
    }
}

/**
 * Recursive function for adding chains of friends in
 * a depth-first traversal.
 */
void addFriendChain(Student s, Activity a, ArrayList<Student>
    participants) {
    if (a.isFull())
```

```
        return;
    a.addParticipant(s);
    participants.remove(s);
    for (Student f : s.friends)
        if (participants.contains(f))
            addFriendChain(f, a, participants);
}
```

As shown in Table 3.6, `Random` results in the lowest amount of friend co-participation, `FriendChain` results in the highest, with `InOrder` in the middle. Due to the positive impact of friends (described in the friend rule), increasing friend co-participation increases both the average interest level and the percentage of youth with an individual interest in STEM. These increases are all statistically significant at 95% confidence.

Table 3.6: Comparison of different methods for matching youth to activity groups. Numbers are the average of 50 simulation runs. All differences between methods are statistically significant at 95% confidence.

Matching Method	Friend Co-participation	Average Interest	Youth with an Individual Interest in STEM
Random	25.8%	0.618	24.1%
InOrder	35.8%	0.624	24.2%
FriendChain	52.9%	0.635	24.9%

3.4 Discussion

The model contains a set of rules developed by subject matter experts and based on literature in the learning sciences. Together, they form a hypothesis about what drives increase or decrease in STEM interest. They include peer and parental participation, passionate and knowledgeable adult leaders, and a high degree of choice in STEM-related activities. The study confirms that this ruleset is sufficient to explain the decline in interest seen in the data.

3.4.1 Calibration

Calibration of a complex model is always challenging. With a high number of adjustable parameters, it is difficult or impossible for a human to assess their effects, particularly when they interact with each other. Evolutionary computation is powerful calibration tool for scenarios like this but it must be deployed carefully. EC is only ever as good as your fitness function. Not only does the fitness function need to distinguish between good and bad solutions, it needs to provide a smooth gradient leading from bad to good solutions.

Goodness-of-fit for evolutionary computation

In our case we needed to compare an entire population of agents to a target population and produce a single metric that encapsulates several dimensions (interest in 4 topics and participation in 20 activities). We considered the chi-square and Anderson-Darling tests, but found the best results with Kolmogorov-Smirnov. The two-sample KS statistic is defined as the largest vertical distance between the empirical cumulative distribution functions (ECDFs) of two populations. It's a value between 0 and 1 that reflects the difference between two distributions. It does a fairly good job of providing the smooth fitness landscape we needed, but with a couple caveats.

First, surveys using a Likert scale produce binned data, and the ECDF of binned data is a step function. Averaging several questions creates more potential values, but that just makes more bins and still results in an ECDF with steps. The problem with comparing step-function ECDFs is that a tiny variation in one of the binned values (e.g., such as that caused by limited floating point precision) causes the corresponding step to shift horizontally. If we were to compare the barely-shifted data to the original, we'd get a KS statistic reflecting the full height of that step. As a result, there's a large change in the fitness function comparing two populations that are nearly identical. Fortunately for us, the model output is not binned, so we avoid this problem. Nonetheless, this should be a consideration for anyone using the KS statistic for evolutionary computation.

Second, when two distributions don't overlap, the KS statistic is 1. This is true whether they are far apart or nearly overlapping. A good fitness function is able to distinguish between terrible and slightly less terrible, but the KS statistic would be no help in these cases. This didn't end up being a problem in our case because our initial population already overlaps with the target distribution and the range of the data is tightly bounded.

A general goodness-of-fit metric that handles both overlapping and non-overlapping distributions could be constructed with the following piecewise function:

$$f(A,B) = \begin{cases} \mathrm{KS}(A,B) & \mathrm{KS}(A,B) < 1 \\ 1 + |\overline{A} - \overline{B}| & \mathrm{KS}(A,B) = 1 \end{cases}$$

Where:

$\mathrm{KS}(A,B)$ is the KS statistic between distributions A and B

Comparison of uncalibrated and calibrated model

Before calibration, the interest levels in the model changed more rapidly and the model was far more volatile. This volatility was exacerbated by the feedback mechanism implicit in the interest threshold. Once above the threshold, agents have more scenarios in which their interest increases and fewer in which it declines. This tended to create a significant bifurcation between those above the threshold and those below, resulting in highly bi-model distribution of interest.

The interest threshold was initially set to 0.5, which seemed reasonable as the midpoint of the interest range. However, as the 6th grade survey data shows, summarized in Figures 3.7, 3.8, 3.9, and 3.10, the vast majority of interest levels are higher than 0.5. In fact, with the interest threshold at the pre-calibration default of 0.5, ~95% of the youth were above the interest threshold in at least one topic. This number is clearly far too high when only 16% of students go on to pursue STEM degrees in college (Maltese and Tai, 2011).

Once calibrated, the model output matches the trajectory of the survey data well enough that the difference between model output and the 7th grade surveys is not statistically significant. This is not particularly difficult since the model was calibrated to match the 7th grade surveys, but extending the simulation for a second year and comparing the model output to the 8th grade surveys data is a meaningful validation test. The results show that the model output matches the 8th grade surveys in three of the four topics, the lone exception being life science which receives a sudden uptick in interest during 8th grade (see Figures 3.7, 3.8, 3.9, and 3.10 and Table 3.4).

This exception highlights an important weakness of this model. It may be the case that life science becomes more interesting in 8th grade because 14 year-olds are naturally more interested in the human body than they were at 12. The rules in our model don't account for developmental changes like this. Similarly, 14 year-olds may not respond to parental encouragement the way they did when they were 12. Future work may try to capture the way these drivers change during early adolescence with rules that consider the age of the agent.

3.4.2 Future Research

Given the lack of ABMs related to interest in STEM, there are many opportunities for further research into this important topic. Here are some suggestions for areas to explore.

Survey Refinements

As mentioned in section 3.4.1, the survey data shows an significant increase in interest in life science between 7th and 8th grades which goes against the trend of the other three topics. This may result from developmental changes that take place between the ages of 12 and 14 which aren't reflected in the rules. Future research could explore rules that take age into consideration.

As detailed as they are, the surveys don't include any questions designed to directly ascertain whether the respondent has a sustained individual interest in a STEM topic. This bit of data would allow a more direct validation of the four-stage model of interest development (Hidi and Renninger, 2006), and would obviate the need for a global interest level threshold. By knowing which side of the threshold each youth is on, we could set an appropriate interest threshold for each youth.

Young people are influenced a great deal by their friends and it would have been useful to have a friend network for this population of youth. There are important privacy concerns involved when asking young people to list their friends, and the Human Subjects Review Board was rightly sensitive to them. Nonetheless, it's worth another attempt to gather this valuable data in an ethical way (for example, Podkul and Sauerteig (2015)).

It's been widely-observed that survey takers tend to bias their answers toward the response they perceive as socially desirable (Fisher, 1993). In the case of this project, that would likely mean over-reporting of their activity levels and possibly their interest as well. We attempted to gauge accurate activity levels by providing a subset of participants with a log book in which they would record all their STEM-related activities. This would have provided not only accurate activity levels for that subset, but also a measure of their initial response error that would could have used to correct for over-reporting in all the surveys. Unfortunately, this was not successful. Future surveys might try using indirect measures to get more accurate responses.

Model Wxtensions

Future work along the lines of this model might benefit from introducing a social influence mechanism (Frank, 1998). For example, if a youth is participating in an activity with a friend, that friend's interest level (high or low) will have an impact on the experience.

Organized activities are known to encourage "initiative, identity exploration and reflection, emotional learning, developing teamwork skills, and forming ties with community members" (Weisner, 2007). If they are also shown to foster interest in the STEM fields, a future version of this model should have an explicit rule to that effect.

The longitudinal data collected by the Synergies team is an incredibly valuable resource for anyone studying interest development or opinion dynamics. It contains more data than we were able to utilize in this model, and the data we did use could be used more fully. For example, we calibrated the model by comparing distributions rather than comparing individuals to themselves in subsequent surveys. One reason for this was that some of the surveys from each year were incomplete, as shown in Table 3.1. Only 70 of the youth completed both the 6th and 7th grade surveys, and only 52 of them also completed the 8th grade survey. In this work, we duplicated the data of the 70 youth who completed both the 6th and 7th grade surveys to produce a population of 140 agents. Future work could use *imputation* methods to replace missing survey data, thus reducing the need for resampling and making within-participant analysis more practical.

3.4.3 Broader Implications for Learning Theory

Existing models of interest development focus on the transition from low situational interest to high individual interest(Hidi and Renninger, 2006). When the transition goes the other direction, from high to low, the process may not be driven by the same mechanisms. More research on the drivers of interest decline is needed.

3.4.4 Policy Implications

The two intervention strategies we studied, information brokering and increased friend co-participation, both yielded positive results. The effect of increased friend co-participation was statistically significant but small in magnitude, with a 2% increase in the percentage of youth with an individual interest in STEM, from 24.1% to 24.9% (see Table 3.6). Brokering potentially has a much more significant impact, with a 42% increase in the percentage of youth with individual interest, from 24% to 34% (as shown in Figure 3.15).

One potential intervention strategy that we weren't able to study fully is topic coordination. Communities might institute a program where they cycle through weekly topics. During that week, youth would have the opportunity to participate in a variety of activities related to the same topic. For example, "radio week" would have an exhibit at the museum on the history of radio communications; an after school activity building a simple radio transmitter; another activity learning morse code and exchanging messages on the radios they build; a lesson in science class on electromagnetic radiation; etc.

3.5 Appendix A

3.5.1 Survey questions

The survey includes several sections. The first section asks how often the respondent participates in the following list of activities:

- Use the public or school library
- Visit OMSI
- Visit the zoo
- Visit the children's museum
- Participate in Scouts or 4-H
- Participate in another kind of afterschool program

- Visit or go camping in a state or national park
- Play a team sport
- Play a sport on my own
- Go to a summer camp
- Hike or spend time outdoors
- Garden or grow plants at home
- Do science kits, experiments, puzzles or stuff like that at home
- Read a book or magazine not for school
- Visit web sites to learn about things you're interested in
- Use a computer to play games at home
- Use a computer to communicate with friends
- Watch a TV program about science, math, or technology
- Build or take things apart or repair things
- Train or take care of pets

The responses were given on a five-point Likert scale:

1. Hardly ever or never
2. A few times a year
3. 1-2 times a month
4. 1-2 times a week
5. Almost every day

The survey responses are summarized in the following charts. Note that in most cases, there's not a large difference between 6th and 7th grade activity levels.

The survey also asked a series of questions to gauge how much the respondent likes "finding out about the following things in or out of school":

- What it is like on other planets and exploring space
- How stars and planets form
- Why clouds, rain, and weather happen

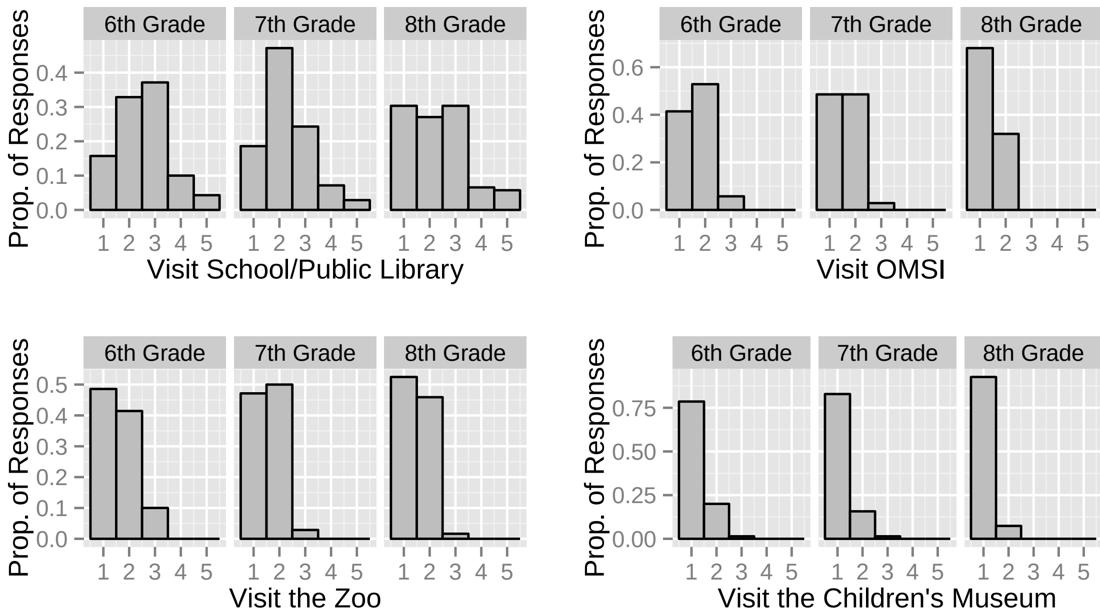


Figure 3.17: Activity Levels reported in the surveys, part 1.

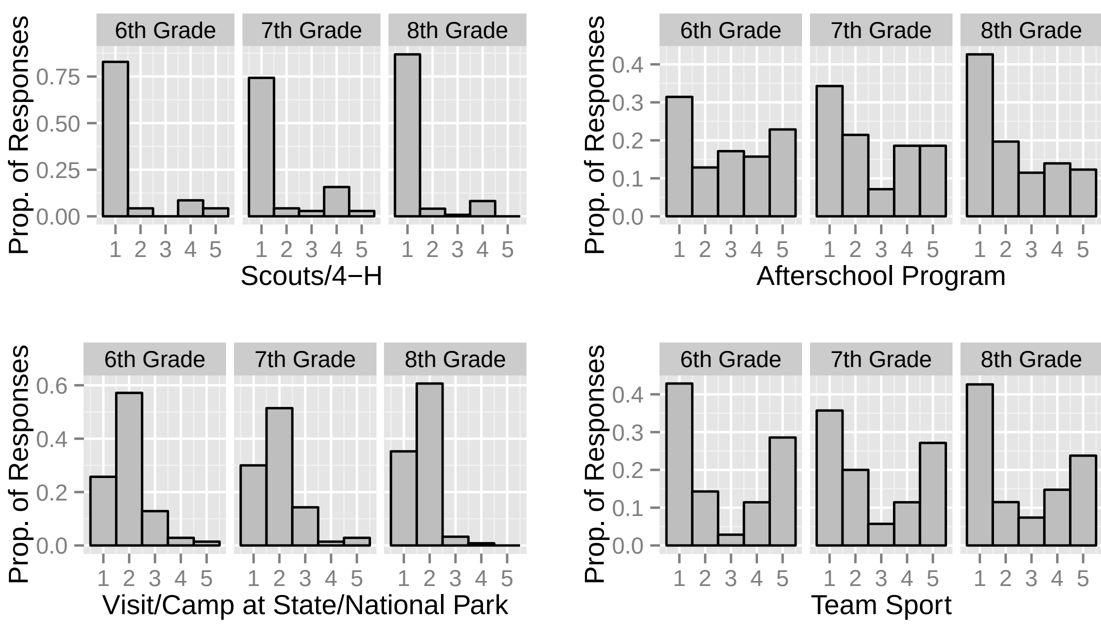


Figure 3.18: Activity Levels reported in the surveys, part 2.

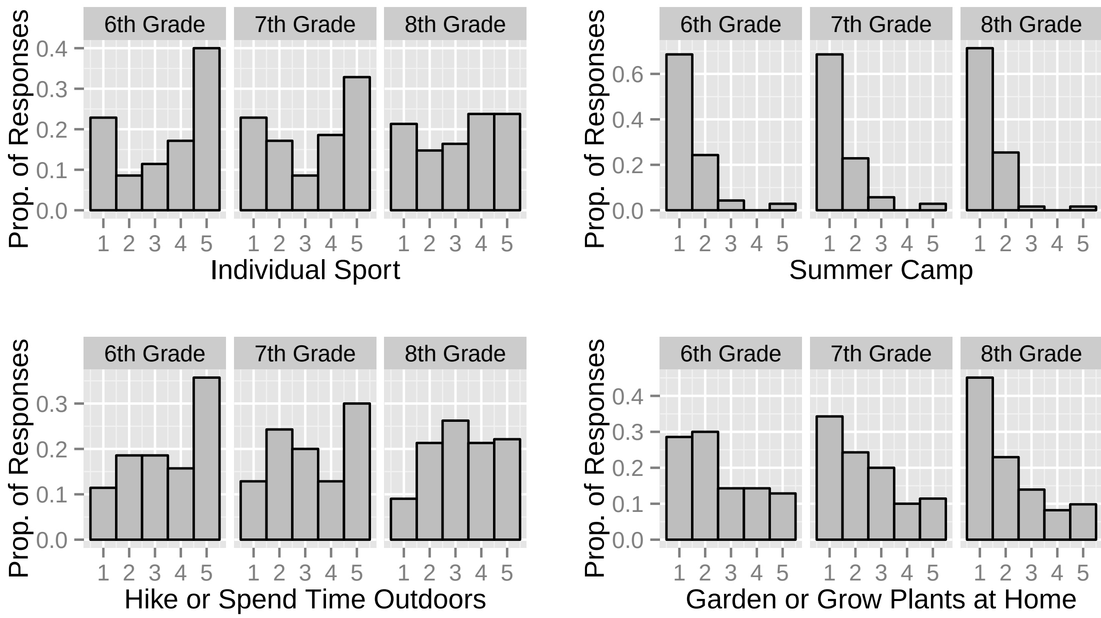


Figure 3.19: Activity Levels reported in the surveys, part 3.

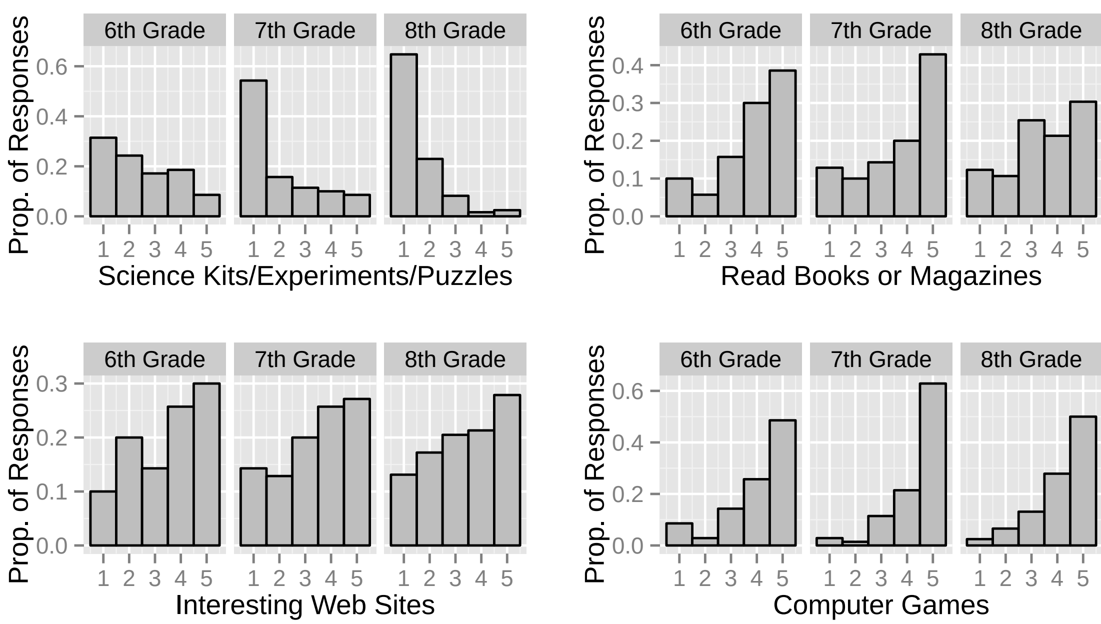


Figure 3.20: Activity Levels reported in the surveys, part 4.

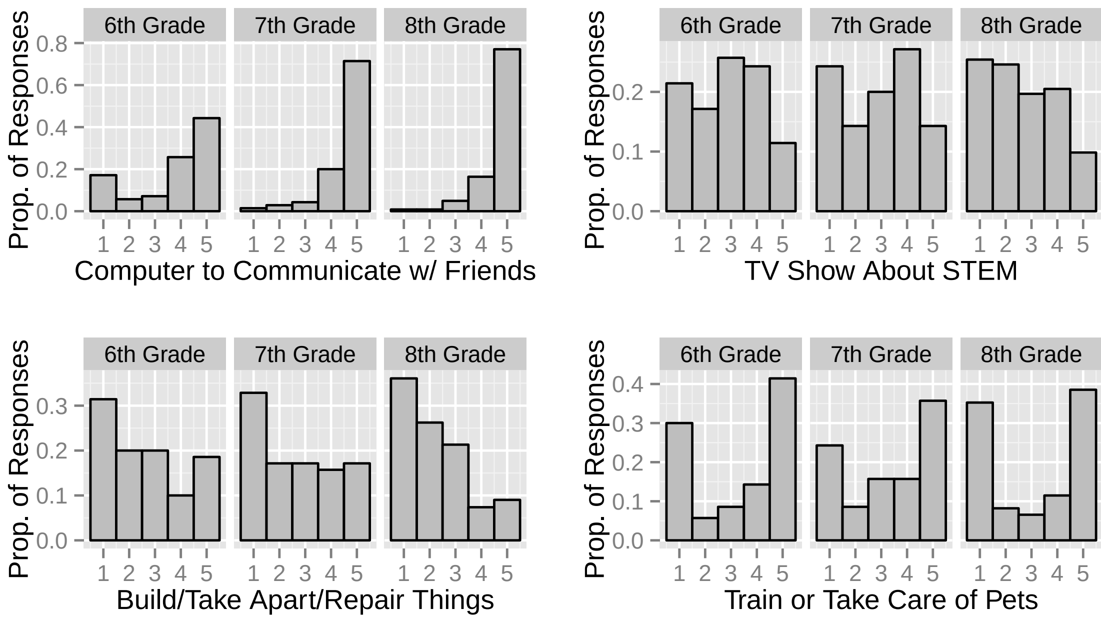


Figure 3.21: Activity Levels reported in the surveys, part 5.

- How earthquakes, volcanoes, and hurricanes happen
- What to eat and how to exercise to keep healthy and fit
- How traits are passed from parents to children
- How the human body works
- How buildings and bridges are made
- How computers, CDs, and cell phones work
- How to use and make maps
- How to design new games or toys
- How gas and diesel engines work
- How to do Sudoku or other math problems
- How to measure the size or area of things
- How to solve puzzles
- How to make different shapes and patterns out of things

The responses to these questions were also given on a five-point Likert scale:

Table 3.7: Principal Component Analysis (PCA) factors for the STEM interest topics. Source: Image from Falk et al. (2015d), Table 1.

Table 1. Summary of rotated factor loadings for STEM interest items in the questionnaire ($n = 249$ 10-/11- and 13-/14-year-old youth, 2012).

How much do you like finding out about[a]:	Earth/space science	Life science	Technology/ engineering	Mathematics
What it is like on other planets and exploring space	0.734			
How stars and planets form	0.811			
Why clouds, rain, and weather happen	0.567			
How earthquakes, volcanoes, and hurricanes happen	0.547			
What to eat and how to exercise to keep healthy and fit		0.701		
How traits are passed from parents to children		0.733		
How the human body works		0.566		
How buildings and bridges are made			0.615	
How computers, CDs, and cell phones work			0.698	
How to use and make maps			0.659	
How to design new games or toys			0.806	
How gas and diesel engines work			0.592	
How to do Sudoku or other math problems				0.749
How to measure the size or area of things				0.670
How to solve puzzles				0.783
How to make different shapes and patterns out of things				0.561
Cronbach's alpha	0.75	0.71	0.78	0.74
% Variance explained	9.84	8.46	40.09	9.67

[a]Items coded on a 5-point scale from 1 = 'Dislike a lot' to 5 = 'Like a lot'. Items in italics were used to construct the STEM knowledge and STEM support indexes.

1. Like a lot
2. Like a little
3. Neither like nor Dislike
4. Dislike a little
5. Dislike a lot

Chapter 4: Radicalization Model and Analysis

4.1 Motivation

In the context social science theory and research, and with major policy concerns arising from the current global environment, it is crucial to understand the radicalization processes that lead to political violence. There are empirically validated theories of individual radicalization (Kruglanski et al., 2009; Sheikh et al., 2016), but it is unclear how these theories play out in social groups. Implementing an agent-based model (ABM) of such a theory would allow exploration of the social process and the emergent phenomena they generate. Once implemented, it would also allow the testing of potential interventions to prevent or reverse radicalization processes

This chapter presents an ABM of individual radicalization based on the work of Kruglanski et al. (2014) and Cioffi-Revilla (2010); Cioffi-Revilla and Harrison (2011). Opinions are modeled using a variation of the continuous bounded confidence model introduced by Deffuant et al. (2000) and expanded by Jager and Amblard (2005) to include a rejection mechanism based on social judgment theory (Sherif and Hovland, 1961).

4.1.1 Terminology

Radicalization is defined by Kruglanski et al. (2014) as "the process of supporting or engaging in activities deemed (by others) as in violation of important social norms (e.g., the killing of civilians)."

Extremism is the holding of extreme political views. The two concepts are closely related with the distinction that radicalization is the process by which a person moves toward extremism.

The opinion dynamics literature talks primarily about extremism. The psychology literature uses both terms, radicalization and extremism, but is centered around radicalization as a process. The opinion dynamics literature has typically referred only to extremism, though some recent work uses both terms (Alizadeh and Cioffi-Revilla, 2014, 2015; Alizadeh et al., 2014, 2015). This paper also uses both terms, along with *extremist*, which refers to a person (or agent) who holds views far outside the political center.

An *interaction topology* is a type of social network through which agents interact. Under a *full-mixing* topology, every agent can interact with every other agent. With *von Neumann* neighborhoods, also know as *4-Set*, agents are embedded in a lattice and only interact with the agents above, below, or to either side. Similarly *Moore* neighborhoods, also know as *8-Set*, allow agents to interact with the eight neighbors surrounding them in the lattice.

4.1.2 Research questions

This work was motivated and guided by the following research questions:

- **RQ4** How do psychological theories (which typically hypothesize intra-individual processes only) play out on the group level? Specifically, how does SQ theory work when implemented with an agent-based model of a simulated community?
- **RQ5** Under what conditions does it create group-level radicalization?
- **RQ6** Do different interaction topologies affect the amount of radicalization?
- **RQ7** What interventions does this ABM suggest to reduce or prevent radicalization?

4.1.3 Background and Prior Work

Significance Quest

The significance quest (SQ) theory of radicalization (Kruglanski et al., 2009, 2014) relies on three components all coming together. First, motivation: people are aggrieved and are motivated to regain personal significance. Second, ideology: binary ideologies provide clarity and meaning, and present violence as an appropriate way to gain significance. Third, a social component: group dynamics and networks which leads to the spread of the ideology and allows carrying out attacks. Dovetailing nicely with Kruglanski, Cioffi-Revilla (2010) provides a formalized theory of radicalization as a multi-path process consisting of three parts: grievance, indoctrination, and the loss of killing inhibition.

The ideological and social component can be implemented in a straight-forward way using opinion dynamics models, which explicitly model the change in opinion based on interactions between agents. At the core of the motivational component of the SQ model is the decreased tolerance of ambiguity and an increased need for clarity: when people are aggrieved they desire clear unambiguous explanations for their grievances, which motivates them to seek out black and white ideologies.

This need for clarity has been successfully assessed with the psychological scale of need for cognitive closure (Webster and Kruglanski, 1994). Individuals with a high NFC seek structure and have a low tolerance for ambiguity (Kruglanski and Webster, 1996). In recent data collections in Morocco (Sheikh et al., 2016), NFC was associated with support for militant jihadism.

Bounded Confidence Models

Deffuant et al. (2000) introduced the bounded confidence model which uses a threshold, u, to represent uncertainty (i.e., the inverse of confidence). Agents interact in pairs and move toward one another if their ideological distance is less than u. The other seminal model in opinion dynamics is Hegselmann and Krause (2002), in which agents are grouped by ideological distance and then all interact together, moving toward the average opinion of the group.

Another early contribution came from Weisbuch et al. (2002), who presented a model based on the Deffuant et al. (2000). They showed that modifying μ or the size of the population doesn't change the outcome, but it does affect convergence times. They also introduced heterogeneous agents with a few being open-minded ($u = 0.4$) and the rest being closed-minded ($u = 0.2$). The open-minded agents move between the two clusters of closed-minded agents gradually pulling them toward each other until they finally reach consensus.

ABMs of Extremism/Radicalization

The opinion dynamics ABMs of radicalization can be categorized broadly into two categories: those with only an acceptance mechanism and those with an explicit rejection mechanism.

Acceptance Only

There are many opinion dynamics models (Deffuant et al., 2002, 2004; Amblard and Deffuant, 2004; Deffuant, 2006) designed to model extremism without an explicit rejection mechanism. These models all begin with a subset of the population designated as extremists, with high confidence (i.e., low uncertainty) in their extreme views. The outcome is then a question of whether the moderates will remain in the middle or be lured to extremism.

Several of these (Amblard and Deffuant, 2004; Weisbuch et al., 2005; Deffuant, 2006) have explored the effects of different interaction topologies with the general conclusion that more restricted interaction results in fewer extremists.

Deffuant (2006) explores the effect of adding noise to opinions during interactions but found that it doesn't significantly change the outcome patterns.

Acceptance and Rejection

The first model to add an explicit rejection mechanism was Jager and Amblard (2005), based on social judgment theory (Sherif and Hovland, 1961). In addition to the ubiquitous uncertainty threshold, u, they added a rejection threshold t. With previous models, agents who are not within u of each other have no influence on one another. This second threshold t represents the distance beyond which another opinion is so repulsive that it causes the agent to move in the opposite direction. Huet et al. (2008) extended this approach to two attitude dimensions. When agents are close in one dimension but far in the other, they'll move away in both dimensions. Alizadeh and Cioffi-Revilla (2014) analyzed the distribution of cluster size in Huet's model and found them fat-tailed.

Salzarulo (2006) built a similar model based on self-categorization theory (Turner et al., 1987) using a principle called meta-contrast, but found little difference compared to the Jager and Amblard (2005) model based on social-judgment theory. This finding was replicated by Crawford et al. (2013).

Exogenous Traumatic Events

Also relevant to this model is the work of Fortunato (2005); Fortunato and Stauffer (2006) who inject extreme events at the start of the simulation which spread through the network directly pushing agents toward extreme opinions.

4.2 Design

4.2.1 Overview

At the core of the SQ theory is the motivational component: after a (traumatic) event people become aggrieved and their need for closure (NFC) increases (i.e., they become less tolerant of ambiguity). The opinion dynamics model of Jager and Amblard (2005), which features an acceptance threshold u and a rejection threshold t, is a good match to model this process because it contains a neutral zone between u and t called the *latitude of non-commitment* (LoNC).

LoNC is a range where agents don't change their opinion after interacting with another agent who is neither close enough, nor far enough, from their own opinion. In other words, the other agent's opinion is ambiguous for them. They're willing to agree to disagree. LoNC provides a straightforward way to implement NFC: NFC goes up, LoNC goes down. When the need for closure is at its highest level (people are highly aggrieved and have no tolerance for ambiguity), LoNC is reduced to zero. An agent must accept or reject, with no room in between: you're either with us, or against us.

When need for closure is low (i.e., person is not aggrieved), LoNC is at the lowest possible value for them. This value, $\text{LoNC}_{\text{calm}}$, differs across agents, (more details on that later).

To inject a source of exogenous grievance, traumatic events are generated as described in subsection 4.6.1. When an agent is hit by an event, it becomes aggrieved and its LoNC reduces somewhat. If an agent is hit repeatedly, it's LoNC continues to shrink until it reaches 0 (i.e., no tolerance for ambiguity).

Over the course of a long model run, many events will be generated, and if their effects continue to accumulate, all agents will eventually become aggrieved. To prevent this, the model has a decay mechanism so that the effect of events eventually wears off and LoNC eventually returns to normal. Even with decay, however, if an agent is hit by several events in rapid succession, the effects will accumulate and reduce LoNC to 0.

Lastly, the Jager-Amblard model Jager and Amblard (2005) has to be adjusted in one important aspect. The original model allows u and t to vary independently of each other[1], and LoNC can be calculated from them[2]. Here, LoNC is set first (by mapping it to NFC) while u and t remain undetermined, and in fact, underdetermined. In the adjusted model, a

[1]with the constraint that $t \geq u$

[2]as simply as $t - u$

change in LoNC changes both parameters (u and t) symmetrically: when LoNC decreases (people become aggrieved and less tolerant for ambiguity), t decreases the same amount as u increases. Changes in u and t are thus perfectly inverted: they mirror each other.

The ideological component can be implemented with a one-dimensional opinion space. This space can represent the range of opinions on any given issue (e.g., Western forces in Muslim countries). Radicalization can be defined as the endorsement of views that are rejected by the larger group because they violate norms like “not killing others”(Kruglanski et al., 2014). Thus, the middle space of the opinion space can be conceived as (at least initially) majority views, while towards the edges (at the extremes) the opinion space represent higher levels of deviation from the middle ground, in other words, higher radicalization. In the interest of simplicity, I focused on a one-dimensional opinion space. This model could be easily extended to examine the dynamics of competing ideologies (e.g., Salafism vs. Wahhabism) using a multidimensional opinion space.

Lastly, the social component of the SQ theory (the interactions with others, social networks, etc.) can be implemented by defining the structure of the network. In this paper, I focused on modeling 4-Set, 8-Set, and full-mixing.

4.2.2 Agents

Attributes

Agents have the following attributes:

Opinion A continuous variable ranging between -1 and 1, with moderates in the middle and extremists on the edges.

LoNC A continuous variable ranging between 0 and 2, with low values corresponding to a high NFC.

$\mathbf{LoNC_{calm}}$ The nominal set point for LoNC when grievance is zero. An agent with a low $\text{LoNC}_{\text{calm}}$ has a high NFC closure even without being aggrieved.

Grievance Current grievance level represented with a non-negative continuous value. Grievance is increased by traumatic events and decays exponentially over time.

Behavior

Agents have a couple of behaviors that are independent of other agents. First, they process traumatic events, increasing grievance in the process. Second, they "heal" over time as their grievance decays exponentially.

Agent interactions

Agents interact in pairs. They compare their opinion to that of the other agent and if they're similar enough, they accept it and move closer on the opinion spectrum. If the other opinion is very far away, they reject it and move away.

4.2.3 Environment

This model uses a very simple physical environment consisting only of a grid. When using a lattice-based interaction topology (e.g., 8-Set or 4-Set), the grid establishes the social network through which agents can interact. This enables the formation of opinion clusters within neighborhoods.

Although the event generator itself has no physical presence in the environment, the events it generates hit agents that have locations. In that sense, the events are part of the environment.

4.2.4 Rules

Inter-agent

When two agents interact, they go through the following steps.

1) Update LoNC based on grievance:

$$\text{LoNC} = \max(0, \text{LoNC}_{\text{calm}} - \text{gImpact} * \text{grievance}) \tag{4.1}$$

2) Calculate temporary u and t thresholds based on LoNC:

$$\begin{aligned} u &= 1 - \text{LoNC}/2 \\ t &= 1 + \text{LoNC}/2 \end{aligned} \tag{4.2}$$

3) Update the agent's opinion based on the same rules as the Jager-Amblard model:

$$\begin{aligned} &\text{if } |x_i - x_j| < u_i \quad x_i = x_i + \mu(x_j - x_i) \\ &\text{if } |x_i - x_j| > t_i \quad x_i = x_i + \mu(x_i - x_j), \end{aligned} \tag{4.3}$$

where agent i is interacting with agent j, and their opinions are x_i and x_j, respectively. Agent i's acceptance threshold is u_i and its rejection threshold is t_i, and μ is the learning rate.

4) If necessary, clip the agent's opinion to keep it within the range of -1 to 1.

Agent-environment

Traumatic events are generated via a Poisson process, the frequency of which is controlled by `eventRate`. This parameter determines the average number of events per day per agent. For example, if there are 100 agents and `eventRate` = 0.1, 10 agents will be hit each day on average. Of course, that's only the average and some days 15 agents will be hit and other days zero.

When an event occurs, an agent is randomly selected as the unlucky recipient. That agent processes the event, increasing its grievance by 1.

4.3 Implementation

The model was written in Java using the MASON simulation toolkit (Luke et al., 2005; Luke, 2015). Using a mature code base like MASON is a tremendous benefit in terms of model verification.

The model was implemented so that parameters can be initialized from the command-line, thus facilitating the use of the ARGO computing cluster for large-scale parameter sweeps and experiments. MASON also has the benefits of being fast and having a small memory footprint, which are both important when trying to run large numbers of simulations.

These factors are important considerations when attempting to answer questions such as **RQ5**: Under what conditions does [this model] create group-level radicalization? In order to know the conditions under which a model exhibits a particular phenomenon, the parameter space must be explored combinatorially. A combinatorial sweep of 6 parameters, each with 20 set points, and 50 repetitions each comes to $20^6 * 50 = 3.2$ billion simulations.

The model was implemented with the agents located in a grid, which supports the use of different interaction topologies such as 8-Set and 4-Set. Full-mixing, by contrast, works equally well in any type of environment or even no environment at all. The ability to choose between different topologies allows us to see if they affect the amount of radicalization (**RQ6**).

4.3.1 Model Parameters

The radicalization model can be configured by adjusting its parameters, which are listed in Table 4.1. During the experiments described in subsection 4.6.2, these parameters are initialized from the command-line.

Table 4.1: Parameters of the LoNC-based radicalization model.

Parameter	Description
width	Width the agent grid
height	Height of the agent grid
mu	Opinion change rate
lonc	Average LoNC assigned to agents
variety	Standard deviation of LoNC assigned to agents
interaction	Interaction topology (full-mixing, 4-Set, 8-Set)
gDecay	Proportion of grievance reduced each step
gImpact	Amount LoNC changes due to grievance
eventRate	Traumatic events per day per agent
stableStepsToStop	Number of consecutive stable steps until the model terminates

4.3.2 Data Preparation

From the citywide Tetouan survey data collected by Sheikh et al. (2016), we calculated LoNC from the questions related to Need for Closure (NFC). Since we are generating events to create grievance, we need to know what LoNC would be in the absence of grievance. We use the survey questions related to Loss of Significance as a proxy for grievance, then calculate NFC_{calm} for each survey participant using regression. Then we invert NFC_{calm}

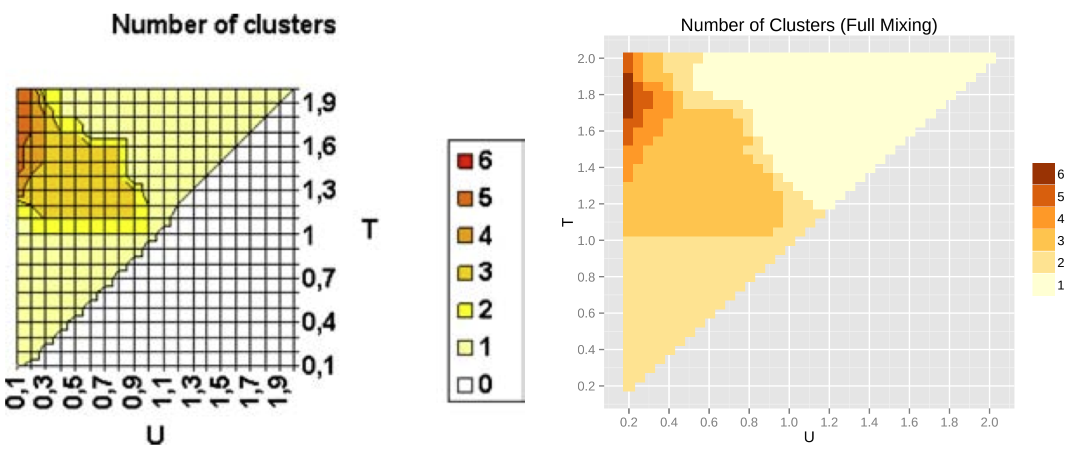
and map it from its range to the range of LoNC. We invert it because a high need for closure indicates a low tolerance for ambiguity, and tolerance for ambiguity is precisely what LoNC represents. The range of NFC is 1–7 and the range of LoNC is 0–2, but the lower half of the LoNC range is not a good starting point. As Figure 5.2 shows, initializing the model with a $\text{LoNC}_{\text{calm}}$ of less than 1 generates at least 50% extremists, and that's without any events. Therefore, we map NFC_{calm} to only the upper half of $\text{LoNC}_{\text{calm}}$, from 1–2. The resulting data has an average $\text{LoNC}_{\text{calm}}$ of 1.59 with a standard deviation of 0.13.

4.4 Verification

This model was verified using code review to examine key components: event generation, event processing, LoNC adjustment, and opinion updating. Each component of the software was also unit tested from the bottom up, verifying that they function properly.

Because the LoNC model is a restricted version of the U & T model, it was possible to compare the preliminary unrestricted model against the results in the original. Figure 4.1 shows the comparison, which is very similar to the original with the exception of lower corner, which turned out to be a printing error in the original paper. This was confirmed by Amblard who provided the original data.

Early sweeps of the Poisson event generator verified the inverse relationship between event rate and events per day, shown in Figure 4.2. The relationship is trivial in hindsight, but at the time I was not expecting to find it. This kind of accidental, blind verification carries considerable weight in establishing model credibility.



(a) Jager-Amblard

(b) Reproduction

Figure 4.1: Comparison of the clusters produced by the original Jager-Amblard and this model. The discrepancy in the lower corner turned out to be an error in original paper. Source: Image from Jager and Amblard (2005).

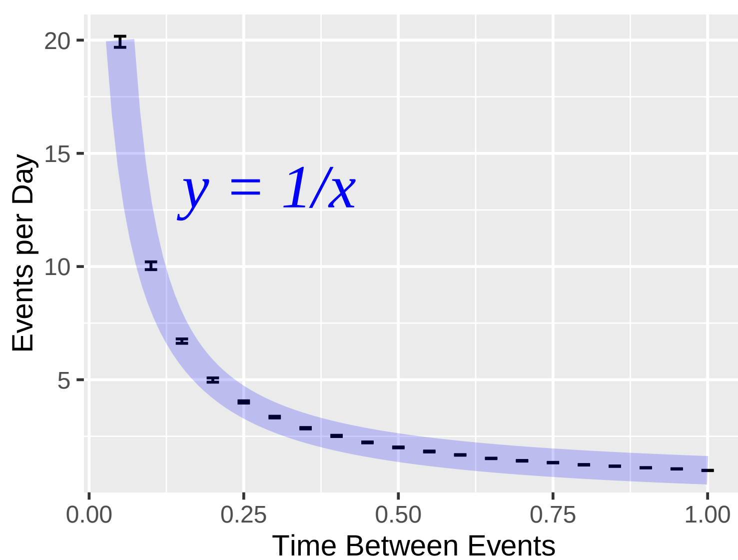


Figure 4.2: Early experiments with the Poisson event generator revealed the unexpected inverse relationship between events per day and time between events. The relationship itself is trivial, but finding it unexpectedly was a strong verification of the event mechanism.

4.5 Validation

4.5.1 Calibration

To ensure that this model works as expected, it's important to find the part of the parameter space where the model is well-behaved. This means two things. First, the function that updates LoNC (shown in Equation 4.1) shouldn't be clipping very often. Clipping occurs when an agent's grievance is so high that LoNC would go negative if it weren't stopped at zero. This happens when `gImpact` is too high, `gDecay` is too low, or there are too many events being generated. If a large percentage of LoNC updates are being clipped, it indicates that additional events are having no effect. Keeping the clipping rate low (<1%) ensures that the model is operating within its effective dynamic range.

Second, the model should produce a fully-radicalized population within a reasonable frequency of events. We assume that if each agent is being hit by an average of one event per step (i.e. `eventRate` =1), that should be sufficient to fully radicalize the population. It may happen at a much lower `eventRate`, but it should definitely happen by the time `eventRate` reaches 1.

Table 4.2 shows the combinations of `gImpact` and `gDecay` for which the model is well-behaved. That is, when `eventRate` is 1, the model clips less than 1% of LoNC updates and radicalization is greater than 95%.

Table 4.2: Combinations of `gImpact` and `gDecay` for which the model is well-behaved over the range of `eventRate` from 0 to 1. I.e., when `eventRate` =1, LoNC clipping <1% and radicalization >95%

	gImpact								
gDecay	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.1	-	-	-	-	-	-	-	-	-
0.2	-	-	-	✓	-	-	-	-	-
0.3	-	-	-	-	✓	-	-	-	-
0.3	-	-	-	-	✓	✓	-	-	-
0.5	-	-	-	-	-	✓	-	-	-
0.6	-	-	-	-	-	✓	✓	-	-
0.7	-	-	-	-	-	-	✓	✓	-
0.8	-	-	-	-	-	-	✓	✓	-
0.9	-	-	-	-	-	-	✓	✓	-

4.6 Analysis

4.6.1 Characteristics of the Event Generator

The combination of cumulative Poisson-generated events and exponential decay results in a metastable equilibrium level which can be calculated with the following equations[3]:

$$\mathbf{E}(g) = \text{eventRate} * \text{eventSize} * \frac{1 - \text{gDecay}}{\text{gDecay}} \tag{4.4}$$

$$\mathbf{E}(\text{var}(g)) = \text{eventRate} * \text{eventSize}^2 * \frac{(1 - \text{gDecay})^{\pi/2}}{2 * \text{gDecay}} \tag{4.5}$$

where $\mathbf{E}(g)$ is the expected value of grievance, and $\mathbf{E}(\text{var}(g))$ is the expected value of the variance of the grievance. Although eventSize has been fixed at 1 throughout the work

[3]Disclaimer: these equations were fit, through trial-and-error, to match experimental data rather than derived from first principles.

for this paper, and can thus be canceled out, the general form of the equation requires its inclusion.

Figure 4.3 shows the shape of this function. The chart is truncated at `gDecay` =0.2 because there's an asymptote where `gDecay` approaches zero.

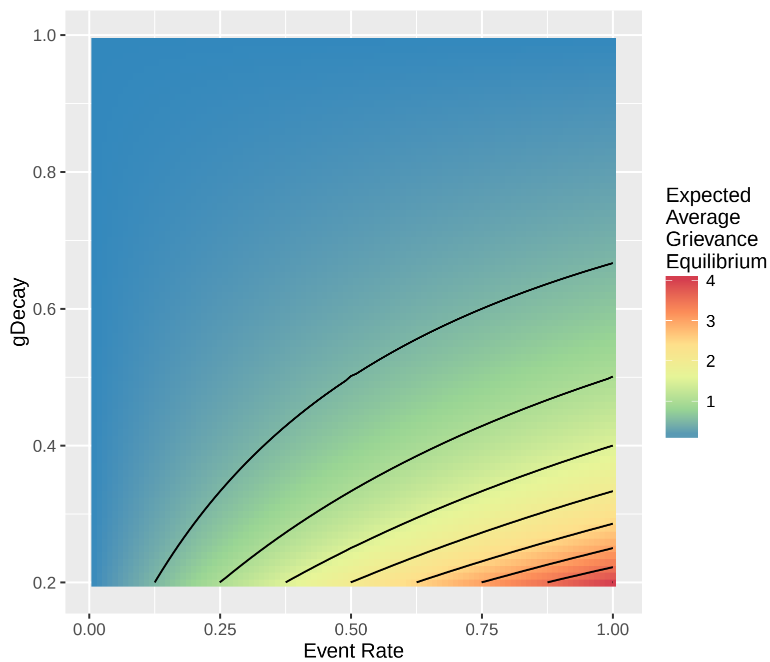


Figure 4.3: Metastable equilibrium levels of the Poisson event generator with exponential decay. Contour lines show grievance levels in increments of 0.5.

4.6.2 Experiments

Each experiment begins by reading model parameters from the command line or from a file. Then the population of agents is created. The agents' starting attributes are either read from a file or sampling from distributions specified in the parameters. For example, the agents in the experiments behind Figure 5.3 were given $\mathrm{LoNC}_{\mathrm{calm}}$ values sampled from a normal distribution with a mean of 1.59 and a standard deviation of 0.13.

The simulation runs until the population converges to a stable state, or for 10,000 steps, whichever comes first. The simulation checks for convergence by counting the number of consecutive steps since the output variable, number of extremists, has changed. If the number of extremists hasn't changed for 250 consecutive steps, it is considered converged. This value is controlled by the command-line parameter `stableStepsToStop`.

When the run stops, it calculates a variety of summary statistics and prints them (along with parameter values). The most important of these, since it's the primary objective of this project, is the count and proportion of extremists.

Agents within 5% of either end of the opinion range are considered extremists. Since opinions range from -1 to 1, this includes agents above 0.9 or below -0.9. The initial opinions of agents are sampled from a uniform distribution within that range, so at step 0 there will already be ≈10% extremists. Even with a LoNC of 2.0 (as open-minded as possible), the model converges into a final state with ≈10% extremists. Thus, we can consider 10% to be the baseline level of extremism this model can generate.

LoNC and Variety without Trauma

To explore the behavior of the new LoNC model, the first experiment varies LoNC from 0–2 and variety from 0–0.5, without the introduction of traumatic events. That is, agents are initialized with LoNCs drawn from a normal distribution with a mean of LoNC and a standard deviation of variety. Throughout each run, agents interact with each other and update their opinions, but their LoNCs don't change because there are no traumatic events. Nonetheless, some portion become extremists based on their starting LoNC alone. The results are shown in subsection 5.1.1.

Lonc, Variety, and Interaction Topology without Trauma

This experiment looks at the role of how variety and interaction topology affect the number of extremists generated by the model, again, without traumatic events. LoNC is swept from 0–2, variety is swept from 0–0.3, and full-mixing, 8-Set, and 4-Set topologies are used. The results are shown in subsection 5.1.2.

Event Rate with the Calibrated Model

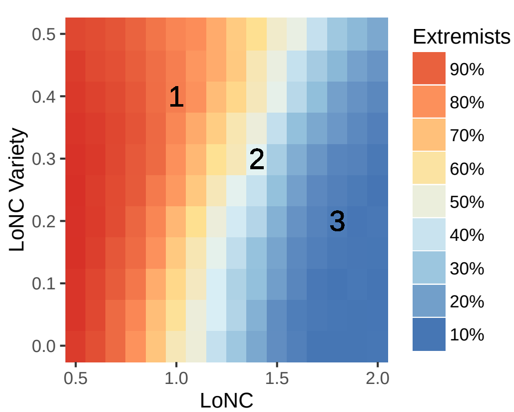
After calibrating to the Tetouan data and identifying the regions parameter space where the model is well-behaved (see Table 4.2), this experiment looks at the response to traumatic events under different interaction topologies. The Tetouan data contains $\text{LoNC}_{\text{calm}}$ with an average of 1.59 and a standard deviation of 0.13. The agents are initialized from this distribution. Selecting from the values in Table 4.2, `gImpact` is 0.2 and `gDecay` is 0.3. The results can be found in subsection 5.1.3.

Chapter 5: Radicalization Model Results and Discussion

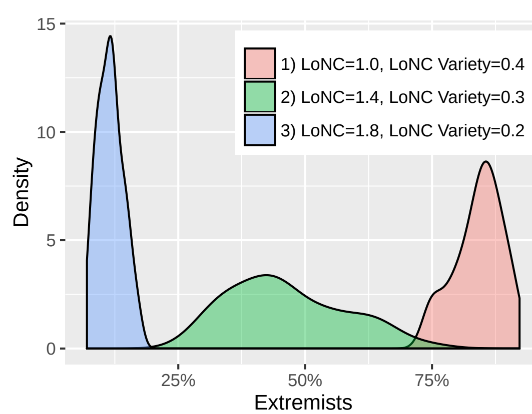
5.1 Results

5.1.1 LoNC and Variety without Trauma

This experiment, described in section 4.6.2, looks at the output of the new LoNC-based model in the absence of traumatic events. Figure 5.1 shows that more extremists result from both lower LoNC and higher variety in LoNC. These runs used the 4-Set interaction topology, i.e., von Neumann neighborhood. At each set point of LoNC and LoNC variety, the model was run 50 times, and the cells in Figure 5.1a show the average of those runs. At the start of each run, the population of agents is initialized with LoNC values drawn from a normal distribution with a mean of LoNC and a standard deviation of LoNC variety.



(a) Average outcome



(b) Distribution of outcomes

Figure 5.1: Lower LoNC results in more extremists. Higher variety of LoNC results in more extremists. The cells labeled 1, 2, and 3 in (a) are detailed in (b), which shows kernel densities for each distribution.

Figure 5.1b shows the distribution of model runs at the three cells labeled 1, 2, and 3 in 5.1a. Cell 1, which has the lowest LoNC and highest LoNC variety, produces the most extremists. Cell 3, which has the highest LoNC and lowest variety, produces the fewest extremists. In the middle, cell 2 produces a medium amount of extremists, but with a high amount of variation in the outcome, with some runs ending with 75% extremists and some only 25%. The range of outcomes at cell 3 is the narrowest, indicating that the model consistently produces less than 20% extremists with those parameters. At cell 1, the model produces the most extremists, but there is more variation in the outcome than there is at cell 3, despite them being equidistant from cell 2 in the parameter space. This is a reflection of the fact that at cell 1, LoNC is in the middle of its range. Moving cell 1 from LoNC=1.0 to 0.8 would reveal a narrower distribution of outcomes more symmetrical to cell 3.

5.1.2 Lonc, Variety, and Interaction Topology without Trauma

This experiment, described in section 4.6.2, shows how the LoNC model behaves, without traumatic events, under different amounts of variety and different interaction topologies (see Figure 5.2).

With no variety in LoNC (top row of Figure 5.2), the general pattern is an s-curve going from ≈10% extremists when LoNC is high to ≈100% when LoNC is low. This is true for all three interaction topologies, though the more restricted topologies (8-Set and 4-Set) produce steeper, rounder curves similar to the standard sigmoid function, while full-mixing produces a more linear curve.

As the amount of LoNC variety increases (going from the top row, downward), the number of extremists increases. This happens gradually for 8-Set and 4-Set, but with full-mixing, there is more complex result. Looking at the chart where variety is 0.2, there are two distinct loci of outcomes. The lower locus is similar to the curve at the top (when variety is 0), while the second locus looks like the curve at the bottom (when variety is 0.3). It appears that increasing variety changes the proportion of runs that end up in each locus.

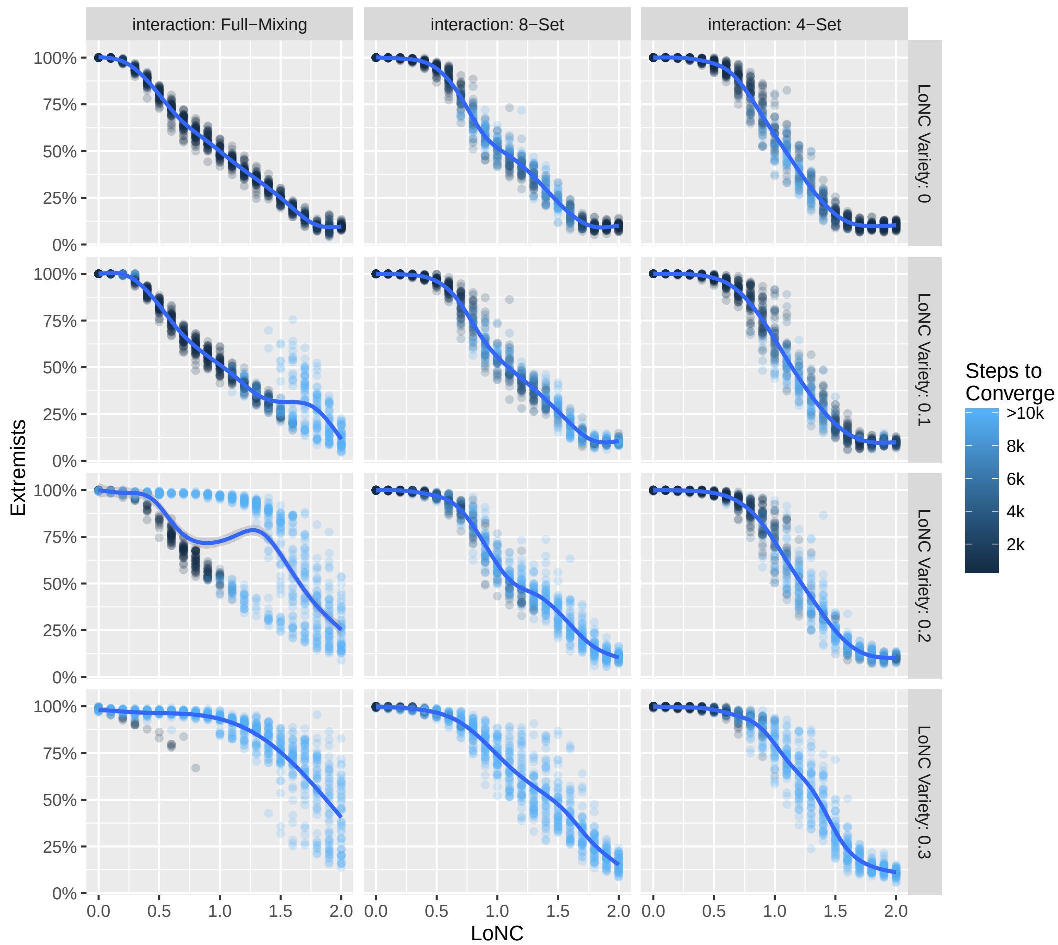


Figure 5.2: As LoNC variety increases, full-mixing eventually leads to more extremism due to increased pathways to the edge. Similarly, 8-set leads to more extremism than 4-set. The curves are drawn by ggplot2's geom_smooth, which uses a generalized additive model to fit the data.

But why should increasing variety in LoNC result in more extremism? The reason might be similar to the finding of Weisbuch et al. (2002) that open-minded agents can move between two clusters of closed-minded agents, and eventually pull the clusters together in consensus. In this case, it would be the higher LoNC agents moving between the edge and the middle, pulling agents to the edge.

5.1.3 Event Rate with the Calibrated Model

This experiment, described in section 4.6.2, uses the model calibrated to the Tetouan data, initializing $\text{LoNC}_{\text{calm}}$ values with an average of 1.59 and a standard deviation (variety) of 0.13. The results, shown in Figure 5.3, demonstrate that as the number of traumatic events increases, the number of extremists increases also. This supports the basic premise of Significance Quest theory that grievance drives radicalization.

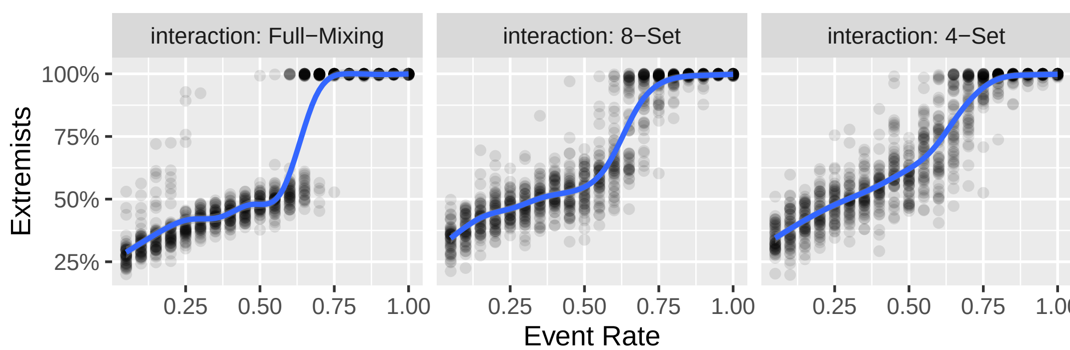


Figure 5.3: Traumatic events increase extremism. The model is calibrated to the Tetouan data where $\text{LoNC}_{\text{calm}}$ =1.59 and LoNC variety=0.13. Picked from the range of well-behaved parameters combinations in Table 4.2, `gImpact` =0.2, `gDecay` =0.3. The curves are drawn by ggplot2's geom_smooth, which uses a generalized additive model to fit the data.

Looking at the shape of the curves, there's a critical point around eventRate=0.6 where the population jumps from 50% extremists to 100%. It's most distinct with full-mixing, but it's also evident under the more restricted lattice topologies, 8-Set and 4-Set. This suggests that it's difficult to remain moderate in a population with a lot of extremists. Consider a hypothetical population where 75% of the people are radicalized and the rest remain moderate. With full-mixing, none of the simulation runs end Figure 5.3 with those proportions (with the exception of a few outliers when eventRate was low). The 75% outcome is slightly more common with 8-Set and more common still with 4-Set. Restricted interaction topologies, it seems, allow moderates to persist in populations with a lot of extremists.

Another interesting difference between interaction topologies is revealed when they are overlapped, as in Figure 5.4. At the low end of the eventRate range, the model produces fewer extremists under full-mixing than with the more restricted interaction topologies. This suggests that exposure to diverse opinions helps to slow the increase of extremism.

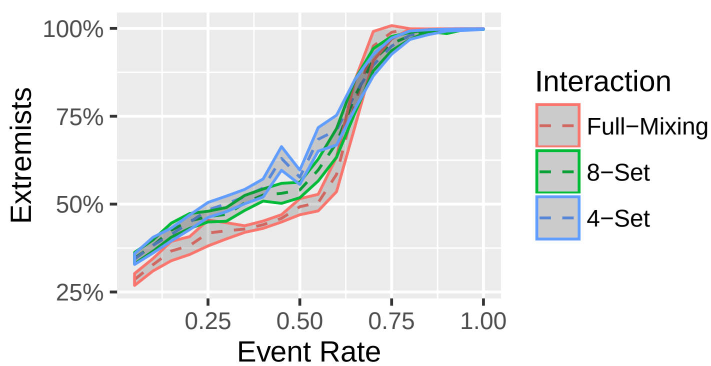


Figure 5.4: Full-mixing has less extremism than 8-Set and 4-Set when there are relatively few events. Bands show the 95% confidence interval.

5.2 Discussion

5.2.1 Research Questions Revisited

Recall the research questions posed in section 1.2 and reiterated in subsection 4.1.2.

RQ4 How do psychological theories (which typically hypothesize intra-individual processes only) play out on the group level? Specifically, how does SQ theory work when implemented with an agent-based model of a simulated community?

The SQ theory predicts that individual grievance (triggered by traumatic events that lead to a loss of significance) translates into radicalization. Figure 5.3 shows that the SQ works as implemented in this ABM. More events generate more extremists. It isn't necessary for the events to modify opinions directly. It is sufficient to merely reduce LoNC, which is to say, increase the need for certainty.

The results of this model show that individual grievances can also lead to group radicalization, providing initial support for the social aspect of SQ theory.

RQ5 Under what conditions does it create group-level radicalization?

The model generates group-level radicalization whenever there are enough events to significantly reduce LoNC. This happens when `gDecay` is sufficiently low enough and `gImpact` is sufficiently high over a given range of `eventRate`. Table 4.2 shows the valid combinations of `gDecay` and `gImpact` for a population of agents calibrated to the Tetouan data. Under the assumption that a community can be fully radicalized with an average of at least one event hitting each agent per step, the model is well-behaved for the combinations of parameters shown in Table 4.2.

RQ6 Do different interaction topologies affect the amount of radicalization?

Yes. Figure 5.2 shows that even without events, different interaction topologies result in different results. The lattice-based networks, 8-Set and 4-Set (Moore and von Neumann, respectively), are only different by degree. With full-mixing though, the difference is drastic when variety is added to LoNC. When variety is above ≈ 0.2, full-mixing produces high levels of extremism even when the average LoNC is high. This is due to increased pathways to the edge, similar to the result found in Weisbuch et al. (2002) where open minded agents were able to move back and forth between clusters and eventually pull them together.

With the model calibrated to data from Tetouan, Morocco, full-mixing produces fewer extremists than 8-Set and 4-Set when events are rare, but more extremists when events are frequent. This result partially contradicts the general finding that restricted interaction leads to lower extremism (Amblard and Deffuant, 2004; Weisbuch et al., 2005; Deffuant, 2006), which is the result of this model when events aren't present but there is increasing variety of LoNC, as shown in Figure 5.2.

RQ7 What interventions does this ABM suggest to reduce or prevent radicalization?

The effect of interaction topology described in the answer to **RQ6** suggests that exposure to diverse opinions helps keep extremism low under normal conditions, but when the amount of extremism goes above a critical point, it may be beneficial to restrict communication networks.

5.2.2 Contributions

For the first time, a social-psychological theory of radicalization that has empirical support with real life extremists (e.g., former Tamil Tigers who participated in the Sri Lankan civil war) has been validated using agent-based modeling. The implementation of a psychological (i.e., individual focused) theory in a model that simulates social interaction between individuals allows us to examine how the theory could affect groups of people, and if it would lead to the hypothesized effects once interactions between people are considered.

In addition, this modeling effort of Significance Quest theory led to a seemingly paradoxical insight that has—to the best of my knowledge—not been explicitly discussed in the literature: while aggrieved individuals are expected to become more closed-minded (that is, they are likely to reject opinions which they tolerated when they were not aggrieved), they also need to become more susceptible to radical ideologies (which they regarded neutrally when they were not aggrieved). So, in a sense grievance leads to both more openness to unambiguous ideologies and less openness to more ambiguous opinions. This is reflected in the implementation as grievance increases both, the willingness to accept other opinions (higher acceptance threshold) and the willingness to reject other opinions (lower rejection threshold).

To implement the core process of the SQ theory—grievance due to (traumatic) events, which motivates the agents to seek out radical ideologies in the first place—a modular random event generator was implemented. This event generator is well-behaved and has only two parameters: the expected rate of the events (`eventRate`) and a decay rate (`gDecay`) which allows the overall grievance (accumulation of events) to wear off over time. This event generator is psychologically and naturally plausible. The events are generated using a Poisson distribution which is commonly observed in events that occur over time (car accidents, bankruptcies, suicides, network failures, etc. (Letkowski, 2012)), while the decay follows a exponential decay (the higher the value the faster it decays per step) which is observed in natural processes such as radioactive decay but more importantly in psychophysiological processes like the decay of perceived pain (Ercole and Roe, 2011), emotional arousal (Feinstein, 2012), and memory over time (Lu et al., 1992). These two parameters lead to an equilibrium point of average grievance, which can be calculated from `eventRate` and `gDecay` as shown in Equation 4.4.

In summary, this is the first agent-based model of radicalization that:

- Is based on an existing social-psychological theory of radicalization (e.g., significance quest)
- Incorporates motivational elements (e.g., desire to regain significance)
- Uses real-world data to any significant degree

In the field of opinion dynamics more broadly, it is the first ABM that:

- Models latitude of non-commitment (LoNC) as a primary variable

- Links LoNC to need for closure
- Uses thresholds drawn from a distribution
- Dynamically modifies thresholds based on exogenous events

Furthermore, this is the first work in any social science—as far as I know—that explores the intermixing of people with varying LoNCs. Given that this study shows increased LoNC variety creating greater extremism, this is of potentially great importance.

5.2.3 Future Reserach

As with many rich projects, this one immediately suggests several opportunities for future research.

Interventions

Now that the well-behaved part of the parameter space has been identified Table 4.2, future research can use this model in that range to test different scenarios and interventions. One such intervention might take inspiration from the STEM Interest model in chapter 3. The study of that model revealed the effectiveness of brokering, wherein a knowledgeable adult directs youth to resources that are relevant to their interests. Brokering is effective at fostering interest in STEM for two reasons. First, it exposes youth to information related to their interests that they may not have seen otherwise. Second, it empowers individuals by giving them control of the activity.

In the context of radicalization, brokering might mean directing individuals to resources that advocate nonviolence or some other peaceful means of conflict resolution. This would be analogous to directly modifying that person's opinion toward the moderate middle of the spectrum. An indirect approach would be to promote mental habits that encourage emotional resilience, such as cognitive behavioral therapy or stoicism. In modeling terms, this could be done by increasing gDecay or reducing gImpact. These are both currently population-wide parameters and would need to be made heterogeneous to explore this approach. Another indirect approach would be to encourage a high LoNC by promoting the non-judgment, detachment, and open-mindedness found in many philosophical and faith traditions. This would be analogous to increasing $\mathrm{LoNC}_{\mathrm{calm}}$, which reflects a decrease in a person's need for closure.

Group Events

The events in this model only befall individuals, causing individual grievance. Another important component may be group events, which affect everyone in the population who belong to the group in question. Such an effect would be moderated by an individual's empathic concern or group narcissism.

Collect More Data and Use More of the Existing Data

This model was calibrated using the Tetouan data but there is more in the dataset that could be used. The current opinion can be taken from an individuals support for violent jihad. Their current grievance level, which this model initializes to zero, could be instead taken from their reported feeling of loss of significance. For the facilitation of group events, an individual's group sensitivity could be taken from the group narcissism data.

Ultimately though, longitudinal data is needed to validate the dynamics of the model, and to explore the true effectiveness of interventions. Future validation studies can collect longitudinal data and check how well the model matches with empirical observations. In particular, collecting data before and after a controlled intervention experiment is the best (and possibly only) way of assessing the true effectiveness of the intervention.

Compare Network Types

Given that we observe different patterns of radicalization for full-mixing and the more restricted lattice topologies, it's worth investigating how the model behaves with more realistic networks such as small-world (Watts and Strogatz, 1998) and scale-free (Albert and Barabási, 2002).

Part III

Discussion and Conclusions

Chapter 6: Discussion

6.1 Compare and Contrast

There are similarities and differences between the STEM Interest and Radicalization models at each stage of the project. This section breaks them down. For the sake of brevity, the STEM Interest and Radicalization models are referred to as STEM and Rad, respectively.

Design

STEM and Rad are both based on the general model described in chapter 2, and they both use real-world survey data. The opinion updating rules for both are based on leading theories of subject matter experts.

By contrast, STEM has data that was collected in conjunction with the modeling effort, which informed the data collection process. Rad was designed to make the best use of existing data, which fortunately was collected by surveys attempting to valid the same QFS theory Rad is based on. STEM uses data-rich agents that are initialized with a real population almost like a micro-simulation, while Rad is more abstract.

STEM's interaction events are activities with one to many participants. There are 20 different activities with different content, but they're all processed under the same rules. Agent attributes determine how often they participate in each activity. Rad's interaction events are pair-wise conversations which modify agents' opinions based on their ideological distance and LoNC. Rad also has traumatic events which modify LoNC, but not opinion. Traumatic events are processed individually and their occurrence is independent of agent attributes.

Implementation

STEM and Rad are both object-oriented and implemented in Java using the MASON toolkit. They are both also designed to accept command-line parameters so they can be run in large-scale experiments on the ARGO computing cluster.

In STEM, the activities are the steppable objects that get scheduled, while in Rad the agents and event generator are steppable. STEM runs for a fixed amount of time (1 or 2 years), while Rad runs until the population stabilizes (or 10k steps if it never does).

Verification and Validation

Both STEM and Rad were verified using code review and unit tests from the bottom up.

By contrast, STEM was calibrated using evolutionary computation with a complex multi-factor fitness function comparing model output to survey data, while Rad was calibrated using parameter sweeps to find the range where the model behaves well (i.e., exhibits the full range of extremism with minimal clipping). Because STEM has longitudinal data, it can be validated by comparing model output to the 8th grade data (after being calibrated to track from 6th to 7th grade). That can't be done with Rad until longitudinal data becomes available showing how grievance (i.e., loss of significance), need for closure, and extremism change over time. Correlating that data in time with a stream of real-world traumatic events would allow for the calibration and validation of the traumatic event generator itself.

Results

Both models identified potentially effective interventions that were not foreseen at the start of the project. In the case of STEM, brokering access to tailored information resources shows promise as cost-effective approach. With Rad, the unexpected intervention is to increase exposure to diverse opinions while extremism is low, but restrict communication when it becomes high.

The main difference between the results of the two models is that STEM produces realistic output that is not statistically significantly different from the validation data. Rad, on the other hand, does not produce realistic final populations. Like all opinion dynamics models, it produces stylized facts that are nevertheless illuminating.

6.2 Implications for the General model

The general model is useful when designing agent-based models of opinion dynamics. The extent of its usefulness depends on how widely it can be applied. If every opinion dynamics model can be described using this framework, it has the potential to become a tool for communication between modelers.

6.3 Broader Implications for Theory and Research

Opinion dynamics has traditionally been confined to a handful of simple, abstract models using no data and producing stylized output. They all start with opinions distributed in a uniform random distribution and end with tight little clusters. Incorporating data into these models isn't a natural fit. If you load a complex distribution of real-world opinion data into your model and it quickly clusters together in small groups that look nothing like the original data, how can you hope to produce realistic output? How could such a model ever be validated empirically?

The STEM Interest model might suggest a path forward. It gets initialized with real-world data and produces realistic output. Perhaps its a combination of having multiple opinion updating rules operating in tandem. The rules consider different criteria and may cancel each other out. The rules also have different weights that get calibrated and validated using longitudinal data. Perhaps these are the keys to building more realistic opinion dynamics models.

6.4 Future Research

In addition to the avenues for future research for STEM described in subsection 3.4.2 and for Rad in subsection 5.2.3, here are some opportunities for more work regarding the general model.

Existing opinion dynamics models could be cataloged according to the general model, which would allow for the creation of a taxonomy of opinion dynamics. This project would also provide an opportunity to test, or at least potentially falsify, the hypothesis that the general model is capable of describing any opinion dynamics model. If counterexamples are found, the general model can be revised. Of course, the usefulness of the general model will be diminished if it becomes too complex.

Chapter 7: Conclusions

7.1 Summary

Two-minute Elevator Story

A person's opinions are a function of their internal attributes, their experiences, and their interactions with others. The general model provides a framework for describing these components and how they are related. It works for domains as different as political radicalization and interest in the STEM fields.

In the case of STEM Interest, I worked together with a team of education researchers to design and refine both a model of interest development and a data collection effort to support it. The data reveal the same trend seen throughout the United States. Between 5th and 8th grade, adolescents in this county undergo a significant decline in interest in the STEM fields. The data, and the model designed alongside it, allowed us to formalize the leading theories of interest development, refine them, and validate them against subsequent surveys. Once the model's credibility was established, I explored several different potential interventions for retaining interest. The two that proved the most promising were increasing friend co-participation and having adults serve as information brokers, guiding youth toward resources that match their interests. The model was even able to show that if the goal is to increase average interest, mentors should spend a little time with everyone. If the goal, however, is to increase the number of likely future STEM majors, it's better to focus efforts on only a few.

To model radicalization, I again worked with subject matter experts, this time to formulate rules for an ABM based on a leading psychological theory of radicalization, significance quest. Agents suffer traumatic events which increase their need for certainty. As a result, they seek simple answers and reject opposing views. Often, this leads/pushes them into extremism. I calibrated the model with data from Tetouan, Morocco and used parameter sweeps to identify the area of the parameter space where the model is well-behaved. Comparing interaction topologies showed that, when traumatic events are rare and a population is relatively calm, extremism can be reduced through the exposure to diverse opinions. However, when events are frequent and grievance is high, radicalization may be reduced by restricting communication networks.

7.2 Contributions

This dissertation makes significant contributions in several areas: the STEM Interest model adds to the state of the art in learning sciences; the radicalization model contributes to the field of opinion dynamics and the study of radicalization in social psychology; and the general model is a useful addition to the field of opinion dynamics. These contributions are enumerated in detail below.

STEM Interest

The STEM Interest model presented in chapter 3 is the first ABM that:

- Models the role and effect of STEM-related activities
- Models interest development using the transition from situational to individual interest
- Uses detailed longitudinal survey data in an opinion dynamics model

Radicalization

The radicalization model presented in chapter 4 is the first ABM that:

- Is based on an existing social-psychological theory of radicalization (e.g., significance quest)

- Incorporates motivational elements (e.g., desire to regain significance)
- Uses real-world data to any significant degree

In the field of opinion dynamics more broadly, it is the first ABM that:

- Models latitude of non-commitment (LoNC) as a primary variable
- Links LoNC to need for closure
- Uses thresholds drawn from a distribution
- Dynamically modifies thresholds based on exogenous events

Furthermore, this is the first work in any social science—as far as I know—that explores the intermixing of people with varying LoNCs. Given that this study shows increased LoNC variety creating greater extremism, this is of potentially great importance.

Computational Social Science

In addition to the domain-specific contributions listed above, this dissertation also makes contributions to the field of computational social science as a whole. The general model of opinion dynamics provides a common framework for designing, implementing, and communicating agent-based models of opinion dynamics. chapter 3 and chapter 4 describe the general model's application to two significantly domains.

The modular event generator in subsection 4.6.1, which combines additive random events with exponential decay, has some remarkable mathematical properties. Given an event rate and decay rate, the system reaches a meta-stable equilibrium state. This makes it potentially useful for modeling any phenomenon with rare events and exponential decay. These could include human phenomena such as the perception of pain, emotional arousal, memory over time, and retirement funds (during retirement), along with natural phenomena such as radioactive decay, chemical reactions, toxicity of pesticides like DDT, certain damped oscillators, and the heat transfer between two objects.

Glossary

Agent An individual or group that makes decisions or carries out actions.

Critical point A point at which a system undergoes a drastic nonlinear change.

Learning rate Often denoted as μ, this parameter controls how quickly opinions change in a model and can be tuned to align a model's dynamics with real-world time frames.

Opinion Dynamics A multi-disciplinary field of inquiry concerned with studying the ways in which a person's opinion is affected by interactions with others.

Phase transition A transition from one state to another. Traditionally used in thermodynamics to describe the changes between solid, liquid, and gas. Used in complexity sciences to describe a point at which a system undergoes a drastic change.

STEM Acronym for Science, Technology, Engineering, and Mathematics.

Bibliography

Abelson, R. P., Aronson, E., McGuire, W. J., Newcombe, T. M., Rosenberg, M. J., and Tannenbaum, P. H., editors (1968). *Theories of cognitive consistency: a sourcebook.* Chicago: Rand-McNally.

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47.

Alizadeh, M. and Cioffi-Revilla, C. (2014). Distributions of opinion and extremist radicalization: Insights from agent-based modeling. In *International Conference on Social Informatics*, pages 348–358. Springer International Publishing.

Alizadeh, M. and Cioffi-Revilla, C. (2015). Activation regimes in opinion dynamics: Comparing asynchronous updating schemes. *Journal of Artificial Societies and Social Simulation*, 18(8).

Alizadeh, M., Cioffi-Revilla, C., and Crooks, A. (2015). The effect of in-group favoritism on the collective behavior of individuals' opinions. *Advances in Complex Systems*, 18(01n02):1550002.

Alizadeh, M., Cioffi-Revilla, C., and Crooks, A. (2016). Generating and analyzing spatial social networks. *Computational and Mathematical Organization Theory*, pages 1–29.

Alizadeh, M., Coman, A., Lewis, M., and Cioffi-Revilla, C. (2014). Intergroup conflict escalation leads to more extremism. *Journal of Artificial Societies and Social Simulation*, 17(4):4.

Allen, T. T. and Davis, N. (2010). A simple agent-based social impact theory model of student STEM selection. In *Simulation Conference (WSC), Proceedings of the 2010 Winter*, pages 278–289. IEEE.

Amblard, F. and Deffuant, G. (2004). The role of network topology on extremism propagation with the relative agreement opinion dynamics. *Physica A: Statistical Mechanics and its Applications*, 343:725–738.

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. *Groups, leadership, and men*, pages 222–236.

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *The Journal of Conflict Resolution*, 41(2):203–226.

Bahr, D. B. and Passerini, E. (1998). Statistical mechanics of opinion formation and collective behavior: Micro-sociology. *The Journal of mathematical sociology*, 23(1):1–27.

Balci, O. (2004). Quality assessment, verification, and validation of modeling and simulation applications. In *Simulation Conference, 2004. Proceedings of the 2004 Winter*, volume 1. Ieee.

Barron, B., Martin, C. K., Takeuchi, L., and Fithian, R. (2009). Parents as learning partners in the development of technological fluency. *International Journal of Learning and Media*, 1(2):55–77.

Calvez, B. and Hutzler, G. (2005). Automatic tuning of agent-based models using genetic algorithms. In *Multi-agent-based simulation VI*, pages 41–57. Springer.

Castellano, C., Vilone, D., and Vespignani, A. (2007). Incomplete ordering of the voter model on small-world networks. *EPL (Europhysics Letters)*, 63(1):153.

Chaiken, S. and Trope, Y. (1999). *Dual-process theories in social psychology*. The Guilford Press.

Ching, D., Santo, R., Hoadley, C., and Peppler, K. (2015). On-ramps, lane changes, detours and destinations: Building connected learning pathways in hive nyc through brokering future learning opportunities.

Cioffi-Revilla, C. (2010). A theory of individual radicalization: Modeling complexity in terrorism and political violence. In *Annual Conference of the Midwest Political Science Association. Chicago, IL.*

Cioffi-Revilla, C. and Harrison, J. F. (2011). Pandemonium in silico: Individual radicalization for agent-based modeling. Paper presented at the Annual Convention of the International Studies Association, Montreal, Quebec, Canada.

Clifford, P. and Sudbury, A. (1973). A model for spatial conflict. *Biometrika*, 60(3):581–588.

Conte, R. and Paolucci, M. (2002). *Reputation in artificial societies: Social beliefs for social order*, volume 6. Kluwer Academic Pub.

Crawford, C., Brooks, L., and Sen, S. (2013). Opposites repel: the effect of incorporating repulsion on opinion dynamics in the bounded confidence model. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 1225–1226. International Foundation for Autonomous Agents and Multiagent Systems.

De Jong, K. A. (2006). *Evolutionary computation: a unified approach*. MIT press.

Deffuant, G. (2006). Comparing extremism propagation patterns in continuous opinion models. *Journal of Artificial Societies and Social Simulation*, 9(3).

Deffuant, G., Amblard, F., and Weisbuch, G. (2004). Modelling group opinion shift to extreme: the smooth bounded confidence model. *arXiv preprint cond-mat/0410199*.

Deffuant, G., Amblard, F., Weisbuch, G., and Faure, T. (2002). How can extremism prevail? a study based on the relative agreement interaction model. *Journal of artificial societies and social simulation*, 5(4).

Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(4):87–98.

Dierking, L., Staus, N., Falk, J. H., Wyld, J., Bailey, D., and Penuel, W. (2015). Synergies case study research: Developing deep understandings of youths' stem learning trajectories. *Museology Quarterly*, 29(1):7–14.

Eagly, A. H. and Chaiken, S. (1993). *The psychology of attitudes.* Harcourt Brace Jovanovich College Publishers.

Ebenezer, J. V. and Zoller, U. (1993). Grade 10 students' perceptions of and attitudes toward science teaching and school science. *Journal of Research in Science Teaching*, 30(2):175–186.

Emerson, R. M. (1962). Power-dependence relations. *American sociological review*, pages 31–41.

Epstein, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation*, 11(4):12.

Epstein, S., Lipson, A., Holstein, C., and Huh, E. (1992). Irrational reactions to negative outcomes: Evidence for two conceptual systems. *Journal of Personality and Social Psychology*, 62(2):328.

Ercole, A. and Roe, P. (2011). Mathematical relationships in anaesthesia and intensive care medicine. *Continuing Education in Anaesthesia, Critical Care & Pain*, 11(2):50–55.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.*, 59:255–278.

Falk, J., Dierking, L., Staus, N., Penuel, W., Wyld, J., and Bailey, D. (2015a). Creating community-wide synergies: Case study of a science education informal-formal collaboration. *Connected Science.*

Falk, J., Dierking, L., Staus, N., Penuel, W., Wyld, J., and Bailey, D. (2015b). Understanding and connecting youth stem interest and participation across the community: The synergies project. *International Journal of Science Education, Part B.*

Falk, J. H. and Dierking, L. D. (2010). The 95 percent solution school is not where most americans learn most of their science. *American Scientist*, 98(6):486–493.

Falk, J. H., Staus, N., Dierking, L., Wyld, J., Bailey, D., and Penuel, W. (2015c). The synergies project: Preliminary results and insights from two years of longitudinal survey research. *Museology Quarterly*, 29(1):15–21.

Falk, J. H., Staus, N., Dierking, L. D., Penuel, W., Wyld, J., and Bailey, D. (2015d). Understanding youth stem interest pathways within a single community: The synergies project. *International Journal of Science Education, Part B*, pages 1–16.

Feinstein, J. S. (2012). Examination of the limbic system's role in emotional experience using a human lesion model.

Festinger, L. (1962). *A theory of cognitive dissonance*, volume 2. Stanford university press.

Festinger, L. and Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, 58(2):203.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of consumer research*, pages 303–315.

Flowerday, T. and Schraw, G. (2003). Effect of choice on cognitive and affective engagement. *The Journal of Educational Research*, 96(4):207–215.

Fortunato, S. (2005). Damage spreading and opinion dynamics on scale-free networks. *Physica A: Statistical Mechanics and its Applications*, 348:683–690.

Fortunato, S. and Stauffer, D. (2006). Computer simulations of opinions and their reactions to extreme events. In *Extreme events in nature and society*, pages 233–257. Springer.

Frank, K. A. (1998). Quantitative methods for studying social context in multilevels and through interpersonal relations. *Review of research in education*, 23:171–216.

French, J. R., Raven, B., et al. (1959). The bases of social power. *Studies in social power*, 150:167.

George, R. (2000). Measuring change in students' attitudes toward science over time: An application of latent variable growth modeling. *Journal of Science Education and Technology*, 9(3):213–225.

George, R. (2006). A cross-domain analysis of change in students' attitudes toward science and attitudes about the utility of science. *International Journal of Science Education*, 28(6):571–589.

George, R. and Kaplan, D. (1998). A structural model of parent and teacher influences on science attitudes of eighth graders: Evidence from nels: 88. *Science Education*, 82(1):93–109.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. Viking Press.

Gigerenzer, G. (2008). *Rationality for mortals: How people cope with uncertainty*. Oxford University Press.

Gigerenzer, G. and Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1(1):107–143.

Gigerenzer, G. and Regier, T. (1996). How do we tell an association from a rule? comment on sloman (1996). *Psychological Bulletin*, 119:23–26.

Gilbert, N. and Troitzsch, K. (2005). *Simulation for the social scientist*. McGraw-Hill Education (UK).

Goldhamer, H. and Shils, E. A. (1939). Types of power and status. *American Journal of Sociology*, pages 171–182.

Hansen, D. M., Larson, R. W., and Dworkin, J. B. (2003). What adolescents learn in organized youth activities: A survey of self-reported developmental experiences. *Journal of research on adolescence*, 13(1):25–55.

Hegselmann, R. and Krause, U. (2002). Opinion dynamics and bounded confidence: models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3).

Hidi, S. and Renninger, K. A. (2006). The four-phase model of interest development. *Educational psychologist*, 41(2):111–127.

Holley, R. A. and Liggett, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *The annals of probability*, pages 643–663.

Huet, S., Deffuant, G., and Jager, W. (2008). A rejection mechanism in 2d bounded confidence provides more conformity. *Advances in Complex Systems*, 11(04):529–549.

Jager, W. and Amblard, F. (2005). Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change. *Computational & Mathematical Organization Theory*, 10(4):295–303.

James, W. (1890). *The Principles of Psychology*. New York: Dover Publications.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American psychologist*, 58(9):697–720.

Keren, G. and Schul, Y. (2009). Two is not always better than one a critical evaluation of two-system theories. *Perspectives on Psychological Science*, 4(6):533–550.

Klemm, K., Eguíluz, V. M., Toral, R., and Miguel, M. S. (2005). Globalization, polarization and cultural drift. *Journal of Economic Dynamics and Control*, 29(1):321–334.

Klemm, K., Eguíluz, V. M., Toral, R., and San Miguel, M. (2003). Nonequilibrium transitions in complex networks: A model of social interaction. *Physical Review E*, 67(2):026120.

Kruglanski, A. W., Bélanger, J. J., Chen, X., Köpetz, C., Pierro, A., and Mannetti, L. (2012). The energetics of motivated cognition: A force-field analysis. *Psychological review*, 119(1):1.

Kruglanski, A. W., Bélanger, J. J., Gelfand, M., Gunaratna, R., Hettiarachchi, M., Reinares, F., Orehek, E., Sasota, J., and Sharvit, K. (2013). Terrorism—a (self) love story: Redirecting the significance quest can end violence. *American Psychologist*, 68(7):559.

Kruglanski, A. W., Chen, X., Dechesne, M., Fishman, S., and Orehek, E. (2009). Fully committed: Suicide bombers' motivation and the quest for personal significance. *Political Psychology*, 30(3):331–357.

Kruglanski, A. W., Erbs, H.-P., Pierro, A., Mannetti, L., and Chun, W. Y. (2006). On parametric continuities in the world of binary either ors. *Psychological Inquiry*, 17(3):153–165.

Kruglanski, A. W., Gelfand, M. J., Bélanger, J. J., Sheveland, A., Hetiarachchi, M., and Gunaratna, R. (2014). The psychology of radicalization and deradicalization: How significance quest impacts violent extremism. *Political Psychology*, 35(S1):69–93.

Kruglanski, A. W. and Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological review*, 118(1):97–109.

Kruglanski, A. W. and Shteynberg, G. (2012). Cognitive consistency as means to an end: How subjective logic affords knowledge. In Gawronski, B. and Strack, F., editors, *Cognitive Consistency: A Fundamental Principle in Social Cognition*, pages 245–266. The Guilford Press.

Kruglanski, A. W. and Webster, D. M. (1996). Motivated closing of the mind: "seizing" and "freezing.". *Psychological review*, 103(2):263.

Latane, B. (1981). The psychology of social impact. *American psychologist*, 36(4):343.

Letkowski, J. (2012). Applications of the poisson probability distribution. In *Proc. Acad. Business Res. Inst. Conf*, pages 1–11.

Lu, Z., Williamson, S., and Kaufman, L. (1992). Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science*, 258:1668–1668.

Luke, S. (2011). *The ECJ Owner's Manual*. Updated regularly. Over 280 Pages. Available at http://cs.gmu.edu/∼eclab/projects/ecj/.

Luke, S. (2015). *Multiagent Simulation and the MASON Library*. George Mason University, 19th edition. Updated regularly. Over 350 Pages. Available at http://cs.gmu.edu/∼eclab/projects/mason/.

Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., and Balan, G. (2005). MASON: A multiagent simulation environment. *Simulation*, 81(7):517.

Luke, S., Panait, L., Balan, G., Paus, S., Skolicki, Z., Popovici, E., Harrison, J., Bassett, J., Hubley, R., and Chircop, A. (2007). ECJ: A Java-based evolutionary computation research system, 2000-2007. *URL http://cs.gmu.edu/∼eclab/projects/ecj/*.

Maltese, A. V. and Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in stem among us students. *Science Education*, 95(5):877–907.

Massey Jr, F. J. (1951). The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78.

National Science Board (2015). Revisiting the stem workforce: a companion to science and engineering indicators 2014. Technical report, National Science Board.

Osborne, J., Simon, S., and Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International journal of science education*, 25(9):1049–1079.

Penuel, W. R., Harrison, J. F., Falk, J. H., Hendrey, M., Staus, N., and Dierking, L. (in review). Using agent-based modeling to gain insight into adolescents' interest development. *Journal of the Learning Sciences*.

Petty, R. and Cacioppo, J. (1986). The elaboration likelihood model of persuasion. *Advances in experimental social psychology*, 19(1):123–205.

Pineda, M., Toral, R., and Hernandez-Garcia, E. (2009). Noisy continuous-opinion dynamics. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(08):P08001.

Podkul, T. and Sauerteig, D. (2015). Pathways and network effect: Understanding social ecologies of connected learning. In *Proceedings of the Annual Meeting of the American Educational Research Association, Chicago, IL*.

Salzarulo, L. (2006). A continuous opinion dynamics model based on the principle of Meta-Contrast. http://jasss.soc.surrey.ac.uk/9/1/13.html.

Sanchez, H. A., Wells, B., and Attridge, J. M. (2009). Using system dynamics to model student interest in science technology, engineering, and mathematics. In *Proceedings of The 27th International Conference of the System Dynamics Society*.

Schmidt, M. S. and Pérez-peña, R. (2015). F.B.I. Treating San Bernardino Attack as Terrorism Case. *The New York Times*.

Sheikh, H., Gómez, Á., and Atran, S. (2016). Empirical evidence for the devoted actor model. *Current Anthropology*, 57(S13):S204–S209.

Sherif, M. and Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change.* New Haven, CT. Yale University Press.

Simpson, R. D. and Steve Oliver, J. (1990). A summary of major influences on attitude toward and achievement in science among adolescent students. *Science education*, 74(1):1–18.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological bulletin*, 119:3–22.

Sood, V., Antal, T., and Redner, S. (2008). Voter models on heterogeneous networks. *Physical Review E*, 77(4):041121.

Stonedahl, F. and Wilensky, U. (2010). Finding forms of flocking: Evolutionary search in abm parameter-spaces. In *Multi-Agent-Based Simulation XI*, pages 61–75. Springer.

Strack, F. and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and social psychology review*, 8(3):220–247.

Suchecki, K., Eguíluz, V. M., and San Miguel, M. (2005). Voter model dynamics in complex networks: role of dimensionality, disorder, and degree distribution. *Physical Review E*, 72(3):036132.

Sznajd-Weron, K. and Sznajd, J. (2000). Opinion evolution in closed community. *International Journal of Modern Physics C*, 11(06):1157–1165.

Talton, E. L. and Simpson, R. D. (1985). Relationships between peer and individual attitudes toward science among adolescent students. *Science Education*, 69(1):19–24.

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

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *nature*, 393(6684):440–442.

Webster, D. M. and Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of personality and social psychology*, 67(6):1049.

Weisbuch, G., Deffuant, G., and Amblard, F. (2005). Persuasion dynamics. *Physica A: Statistical Mechanics and its Applications*, 353:555–575.

Weisbuch, G., Deffuant, G., Amblard, F., and Nadal, J.-P. (2002). Meet, discuss, and segregate! *Complexity*, 7(3):55–63.

Weisner, T. S. (2007). Well being, chaos, and culture: Sustaining a meaningful daily routine. *Ecology.*

Xie, H., Cairns, R. B., and Cairns, B. D. (1999). Social networks and configurations in inner-city schools: Aggression, popularity, and implications for students with ebd. *Journal of Emotional and Behavioral Disorders*, 7(3):147–155.

Curriculum Vitae

Joseph F. Harrison received his Bachelor of Science degree in Computer Science from Utah State University in 2002. After working in the robotics industry for several years, he went back and earned his Master of Science degree in Computer Science from George Mason University in 2009, graduating with distinction. After receiving his PhD in Computational Social Science from George Mason University in 2016, he will work as a computational social scientist at SoSACorp.