

# Individual vs. Collective

Ryan Kleinberg, James Lopez, Dr. Wayne Iba

December 9, 2018

## Abstract

...

## 1 Introduction

The purpose of this research was to observe how a minority character trait evolves over time within a population. We wanted to see how cultural norms would be effected by a minority character trait, which would be generally oppressed by the majority. A simple application would be racism in modern societies. These racist members are portraying a character trait that is widely perceived as unwanted and can drive other people to aggression, if provoked. We wanted to see how, or if, such a character trait could survive in preexisting possible hostile circumstances.

## 2 Methods

We initialized a population made up of either good or evil agents. After every iteration, each agent would be paired randomly with another. At this point each agent would decide whether to express their dominate character trait, express nothing, or express their lesser character trait. This means that despite an agent being good it could still have the ability to express as evil. Depending on how an agents partner expressed, the bias that decides how evil/good the original agent was will be updated. Agents who expressed the same will always reinforce their bias toward what they just expressed. Agents who were in an interaction that involved differing expressions compare a random number with their commitment levels. Depending on where the random number is, the agent will either move their bias of expression toward their opposing partner or further away, decreasing or increasing their commitment levels correspondingly.

### 2.1 Agents

An agent is a representation of a person. We want to be able to have a population of people, or agents, that interact with each other. Each agent has two features: an expression bias, and a commitment bias

### 2.1.1 Expression Bias

The expression parameter on agents acts as how often or likely an agent is to express an undesirable trait. The bias is treated as a point on a scale from 0-1. If a random number between 0 and 1 is below the bias an agent will express a negative trait, while if it is above the bias it will express a positive trait. This allows 0.5 to act as a mediator between good and evil agents. The closer to 0.5, the more impartial the agent will be.

### 2.1.2 Commitment

Commitment is a parameter on agents that portrays how committed an agent is to their expressive stance. Commitment controls how much an agent is affected from a negative interaction as well as helping determine what the agent will express.

## 2.2 An Iteration

In a single Iteration, an entire population of Agents is randomly shuffled. Then each agent will have an interaction with a partner. Next, the agent will respond to its partner according to what happened in the interaction. Interactions are further explained in a later section.

## 2.3 Expressions

Expressions are based off each agent's bias. The first thing that is done is a neutral zone is calculated for the agent. The neutral zone is calculated by the following function of bias and commitment:

Right border of neutral zone:  $((1 - \text{Commitment})/3) * (1 - \text{Bias})$   
Left border of neutral zone:  $((1 - \text{Commitment})/3) * (\text{Bias})$

This ensures a neutral zone takes a percentage, based off the commitment, from either side of the bias. The less committed an agent, the larger the neutral zone.

After this, think of a scale from zero to one with a neutral zone surrounding the bias on the scale. A random number is generated between 0 and 1. Based off where the value is on the scale, an appropriate trait is expressed: Less than the neutral zone: -1 (Evil trait) In the neutral zone: 0 (No expression) Greater than the neutral zone: 1 (Good trait). *Note: an agent with Bias of 0.5 and commitment of 0, has an equal chance of expressing good, evil, and nothing.*

## 2.4 Interactions

Interactions are the types of scenarios that can take place when two agents express to each other. For example, if there are two agents that are chosen to

interact with each other, one can express an evil feature while the other can express a good feature. Below is a table of all expression scenarios.

Expression Scenarios
Good vs. Evil
Good vs. Good
Evil vs. Evil
Neutral vs. Good
Neutral vs. Evil
Neutral vs Neutral

Affective Response to Partner Expressions	Description
Positive	The agent has taken the opposing "opinion" constructively. Commitment is shifted toward the opposing Agent's bias. This is supposed to resemble a person's willingness to have an open mind and take in all opinions
Negative	The agent has taken the opposing "opinion" as offensive. Commitment is shifted away from the opposing Agent's bias. This is supposed to resemble a person "doubling down" on their opinion.
Neutral	This happens when an agent's partner does not express anything, or the <i>express</i> function returns a 0. There is no change.

*Note:* If both agents express the same magnitude (-1 and -1, 1 and 1) they will automatically have a positive response

*Note:* If both agents express opposing outcomes (-1 1, -1 0, 0 1) they each generate a random number between 0 1, and when compared to each agent's commitment, their separate responses to their neighbor will be determined:

## 2.5 Updates

The purpose of the `updateAgent` function is to update an agent's bias and commitment based off it's interaction with a partner. There are four parameters involved in updating an agent, and they are laid out in the following chart:

Parameter Name	Description
Agent a	The agent that will be updated
int partnerExpress	The expression that the agent's partner expressed. This is the expression that Agent a is reacting to. Can be -1 (evil), 0 (no expression/neutral), 1 (good)
int myExpress	The expression that the Agent a expressed for that round. Can be -1 (evil), 0 (no expression/neutral), 1 (good).
int quality	This is the quality of the interaction with the partner agent. This will decide whether Agent a reacts positively or negatively to its partner. Can be -1 (negative interaction), 1 (positive interaction)

Agent a's bias is changed according to the following formula:

---


$$\text{Bias} = \text{Bias} - [(1 - \text{commitment}) / 10] (\text{Partner Expression}) (\text{Quality of interaction})$$


---

and the formula for updating Agent a's commitment is:

---


$$\text{Commitment} = \text{Commitment} + [(1 - \text{Commitment}) / 10] (\text{Partner Expression}) (\text{My Expression}) (\text{quality of interaction})$$


---

Both the commitment and bias are fixed within the range of 0 to 1.

## 3 Experiments

### 3.1 Hypotheses

I made multiple hypothesis concerning the different parameter changes in each agent:

#### 3.1.1 Total agents changing their bias (moving over/under .5) per iteration in an experiment

There will start out with a significant amount of changes over the bias. Then there will be a steady decrease in changes, and eventually all agents will solidify their bias standing as the experiment lasts longer

#### 3.1.2 Agents changing their bias from 'evil' to 'good'

Initially, the number of agents changing to evil and agents changing to good will seem the same. After a few iterations, however, we should see that there are much more evil agents changing to good because of the greater amount of goods in the population.

### **3.1.3 Overall changes in commitment and bias**

Every agent's commitment will end very close to 1, in other words every agent will be very fortified in their opinion. Every agent's bias will end very close to 1 or 0, this goes along with agents fortifying their ideologies as the experiment lasts longer.

### **3.1.4 The total amount of expressions over the entire population**

As each agent solidifies its ideology, there will be an increasing number of expressions. Expressions over the entire population will increase as the experiment lasts longer.

### **3.1.5 The total number of harmonious interactions**

The total number of harmonious interactions will increase as the majority population increases in number.

### **3.1.6 Number of differing interactions (when a 'good' and 'evil' expressions both appear in the same interaction)**

These differing interactions will decrease as the majority population increases in size.

### **3.1.7 'Good' vs 'Evil' population**

The majority population will always grow, and depending on preset parameters could have a chance at completely eradicating the minority. A large bump is expected in the majority population size, such that the minority size will often be halved.

### **3.1.8 Total number of 'Good' and 'Evil' expressions**

As the majority population increases so will the number of good expressions. As the number of evil agents decreases in the population, so too will the number of evil expressions.

### **3.1.9 Total number of negative vs positive responses**

The number of positive responses will increase over time and the number of negative responses will decrease.

### **3.1.10 What will happen when the minority has a higher starting commitment level**

With a minority that has a higher starting commitment, they will see an increase in their overall population numbers. The magnitude of the increase will be dependent on the starting population size of the minority. The larger the

minority group within the starting population, the more growth that minority will see.

### **3.2 Experiments**

Each experiment was set up to have a variable amount of agents in the population. The population would be initialized as a good/evil split. Next, the population would enter action cycles or iterations. In each iteration, every agent would pair up with another random agent in the population. Each agent would then express as good, evil, or not express based off the agent's expression bias and commitment. The agent would then react to its neighbor's expression. If the agent experiences an expression that is opposite to its bias it will choose whether to react positively or negatively based off its commitment level. If the agent chooses to react negatively it will solidify its current standing and strengthen its bias and commitment. If the agent chose to react positively to its opposite expressing neighbor, it would weaken its current standing by decreasing its bias and commitment. When an agent and its neighbor express the same characteristic, both will automatically react positively. When an agent's neighbor does not express, then there is no change in the agent's commitment or bias.

### **3.3 Settings**

Every experiment run has five primary, controllable parameters:

#### **3.3.1 Percent Evil**

This sets what percentage of our population will be evil

#### **3.3.2 Number of Iterations**

This determines how many iterations the population will be subject to during the entire experiment. In other words it represents how many times agents would interact with another random agent within the population

#### **3.3.3 Commitment Min/Max**

These parameters set the range of where every agent's commitment levels are to be initialized. These control the starting range of initialization for both good and evil agents.

### **3.4 Bias Initialization**

An agent's bias is initialized differently depending on whether it will be an evil or good agent. For evil agents, their biases are initialized by choosing a random number between .5 and 1. For good agents a random number between 0 and .5

is chosen as the bias. As time goes on these biases will gravitate toward 0 or 1 depending on the interactions of the agent.

## 4 Results

### 4.1 Addressing the Hypotheses

#### 4.1.1 Total agents changing their bias (moving over/under .5) per iteration in an experiment

Our hypothesis is confirmed when we look at figure 7. According to this figure, we see an initial rise in total changes, reaching a peak around iteration 10. Then a sharp drop in total changes which is almost at zero by iteration 80.

#### 4.1.2 Agents changing their bias from 'evil' to 'good'

Inspecting figure 1 we can see that there always seems to be a greater number of agents changing to good compared with agents changing to evil. This contradicts our hypothesis that they initially will match each other. An explanation for this behavior could be that agents who have a bias much closer to 0.5 are more likely to interact with a good agent, giving it a higher chance to move across the 0.5 threshold.

#### 4.1.3 Overall changes in commitment and bias

#### 4.1.4 The total amount of expressions over the entire population

Figure 6 confirms our hypothesis that expressions will increase over time. Analyzing figure 6 will show that expressions start at around 86% of the population, eventually increasing to 100% of the population.

#### 4.1.5 The total number of harmonious interactions

Our hypothesis is confirmed when looking at figure 5. We can see the number of harmonious interactions rise from the low 200's to just below 400 by the end of the experiment.

#### 4.1.6 Number of differing interactions (when a 'good' and 'evil' expressions both appear in the same interaction)

Once again, when looking at figure 5, we can see that the number of differing interactions decreases over time, dropping from around 300 to 100 by the end of the experiment.

#### 4.1.7 'Good' vs 'Evil' population

*The majority population will always grow, and depending on preset parameters could have a chance at comple*

#### **4.1.8 Total number of 'Good' and 'Evil' expressions**

This hypothesis is dependent upon some factors of initialization. The majority population must have equal values for commitment initialization or have at least a 65% majority to begin the experiment. If initialized with weaker commitment levels and a below 65% starting majority, our hypothesis will not holdup. But we can see in figure 3 that our hypothesis is confirmed. The number of good expressions rises while evil expressions fall throughout the experiment.

#### **4.1.9 Total number of negative vs positive responses**

Looking at figure 2 we can see that our hypothesis is confirmed and positive responses increase while the negative decreases. This is most likely due to the fact that in the beginning of the experiment there are more evil agents, giving more opportunity for a good and evil agent to have negative responses. As evil agents decrease there is a greater likelihood that two good agents can have a positive response.

#### **4.1.10 What will happen when the minority has a higher starting commitment level**

Looking at figure 9 we can see different scenarios where the minority population has a higher range of starting commitment. Here the evil agents start out with 50%-70% of the starting population. The evil population is only able to keep their majority when initialized at or above 62%, and the evils are not able to grow unless they are initialized at or above 68%.

These results affirm our hypothesis that when a minority is initialized with a higher commitment level compared to the majority they can see significant growth.

### **4.2 Puzzling Trend**

In an attempt to gain insight on how much differing commitment levels effected outcomes an environment was set up to reflect four different scenarios: Weakly and Strongly initialized commitments for good agents vs. weakly and strongly initialized commitments for evil agents.

The conditions for this experiment were that the range for weakly initialized commitments was 0.0-0.5 and the range for strong commitments was 0.5-1.0. Every time the evils were initialized at 30% of the population, which would be 300 evil agents. The total population was 1,000 agents and each experiment was run with 200 iterations. We can see the results in Chart 1, which are averaged over 50 experiments.

The puzzling trend that shows itself here is that the strong majority of goods fails to eliminate as many weak evil agents when compared to the weak majority of goods. This experiment was run multiple times to double check these results and in each case the weakly initialized good majority converted more evil agents than the strongly initialized good majority. Even in the case of strong against



	Weak Evil	Strong Evil
Weak Good	74	283
Strong Good	78	281

Table 1: These numbers reflect how many evil agents are present at the end of the experiment. There started out with 300 evil agents present. Each result is according to the scenario its row and column titles suggest, for example when there was a weakly initialized evil population with a weakly initialized good population the result was 74 remaining evil agents.

strong there is an insignificant gap between how many evil agents remain. We cannot come up with a plausible explanation for this phenomena.

### 4.3 Experiments

Each experiment was set up to have a variable amount of agents in the population. The population would be initialized as a good/evil split. Next, the population would enter action cycles or iterations. In each iteration, every agent would pair up with another random agent in the population. Each agent would then express as good, evil, or not express based off the agent’s expression bias and commitment. The agent would then react to its neighbor’s expression. If the agent experiences an expression that is opposite to its bias it will choose whether to react positively or negatively based off its commitment level. If the agent chooses to react negatively it will solidify its current standing and strengthen its bias and commitment. If the agent chose to react positively to its opposite expressing neighbor, it would weaken its current standing by decreasing its bias and commitment. When an agent and its neighbor express the same characteristic, both will automatically react positively. When an agent’s neighbor does not express, then there is no change in the agent’s commitment or bias.

### 4.4 Settings

Every experiment run has five primary, controllable parameters:

#### 4.4.1 Percent Evil

This sets what percentage of our population will be evil

#### 4.4.2 Number of Iterations

This determines how many iterations the population will be subject to during the entire experiment. In other words it represents how many times agents would interact with another random agent within the population

#### 4.4.3 Commitment Min/Max

These parameters set the range of where every agent's commitment levels are to be initialized. These control the starting range of initialization for both good and evil agents.

## 5 Analysis

The following subsections are mathematical representations of various aspects of our experimentation. They have been tested and properly reflect what could happen in a simulation.

### 5.1 Expected Value

$$(1 - Nr) - Nl$$

This equation represents what value we can expect a specific agent to express.  $Nr$  and  $Nl$  represent the right and left side of the agent's neutral zone (Neutral zones are addressed in section 2.3). The output should be on a range of -1.0 to 1.0, -1 representing evil, 0 representing neutral or no expression, 1 representing good.

### 5.2 Probability of Expressing

$$(1 - Nr) + Nl \Rightarrow 1 - (Nr - Nl)$$

This equation returns the probability that an agent will express. If an agent expresses 0, it represents them not expressing anything of significance toward their partner. The lowest probability of expression for an agent will be 66% due to the nature of how the neutral zone is calculated.

## 6 Future Work

There are a few things that I think could be an interesting way to expand on the model I have created. First, an implementation of social networking could produce interesting results. Agents would form networks as they have positive interactions. Each agent would have a high chance of drawing a "neighbor" from their network instead of finding a random agent outside their network. I would assume networks would all have the same expression bias and we would see good and evil networks form.

A second implementation is having the agents produce offspring and have a life span. As agents died and pass on their ideologies we could truly see how long the 'evil' characteristic would resist being completely eliminated, if that is a possibility. Figure 1 is a reference



Figure 1: Total Good and Evil Changes by iteration

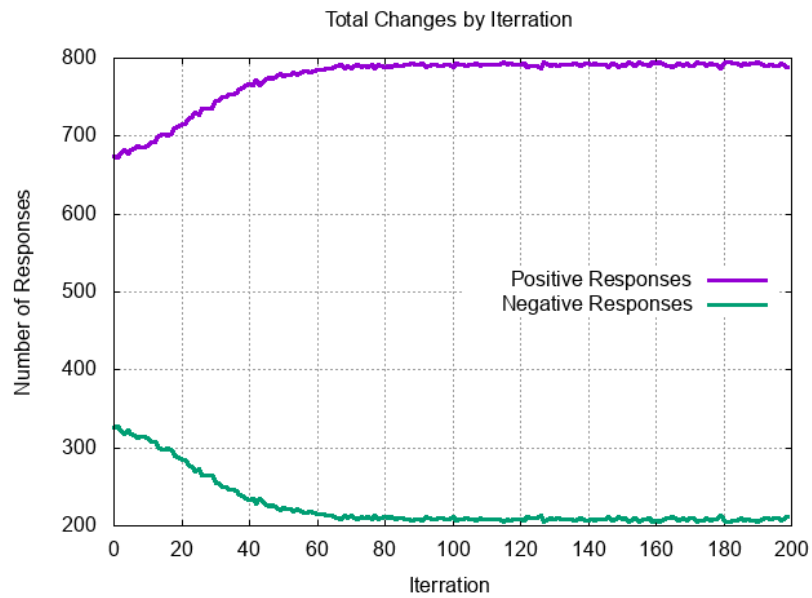


Figure 2: Positive Responses vs. Negative Responses

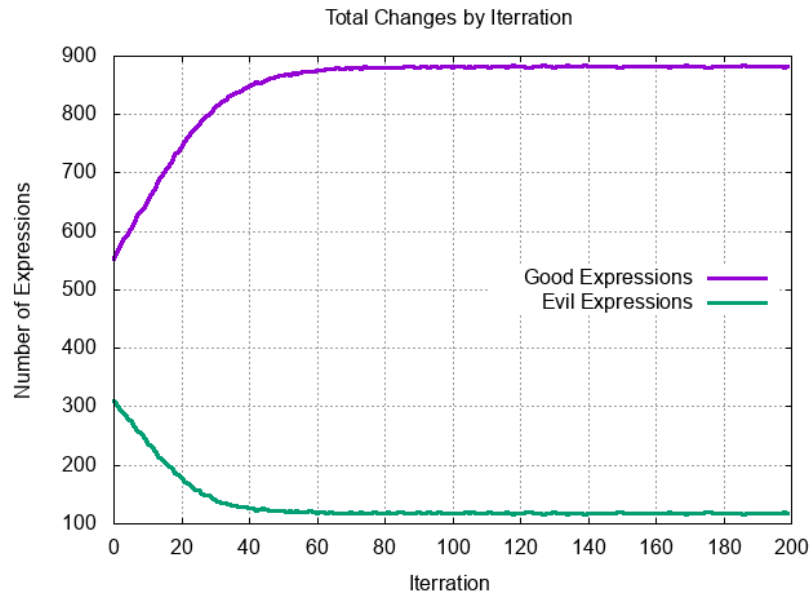


Figure 3: Good and Evil Expressions Over Time

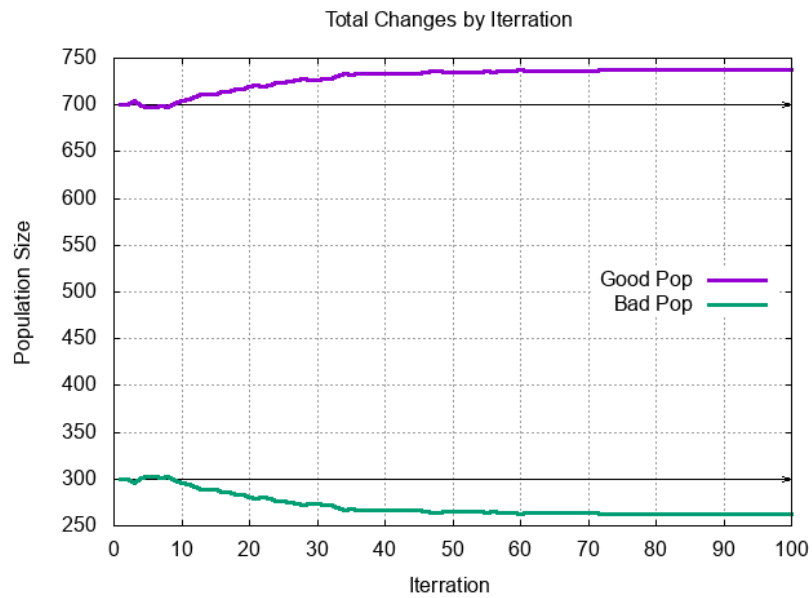


Figure 4: Good and Evil Population Sizes

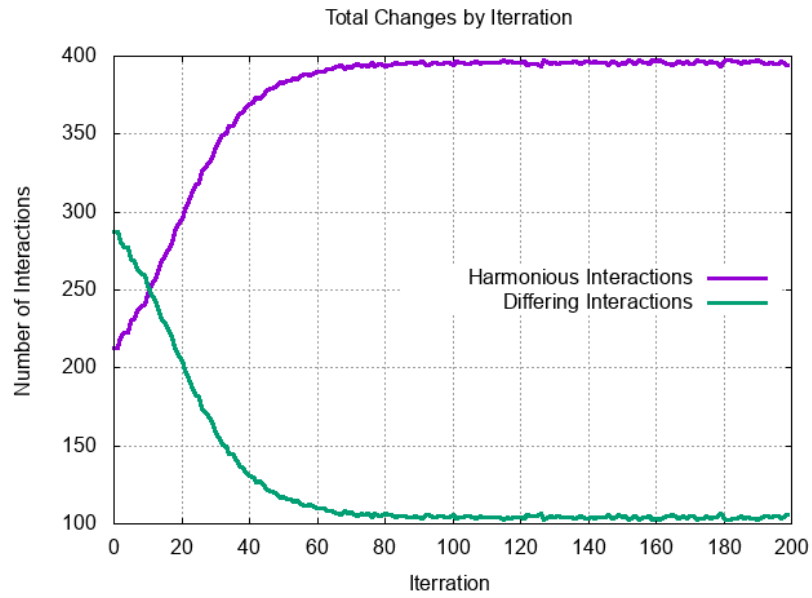


Figure 5: Harmonious and Differing Interactions

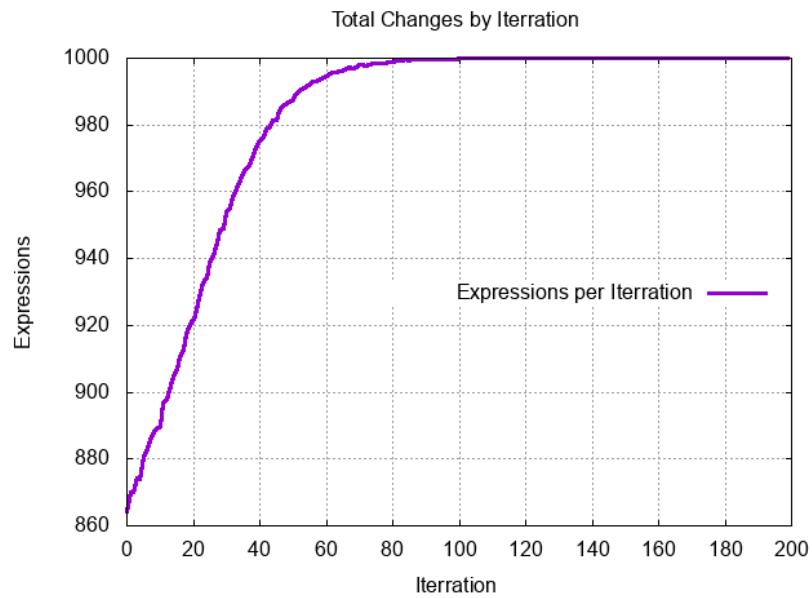


Figure 6: Number of Expressions

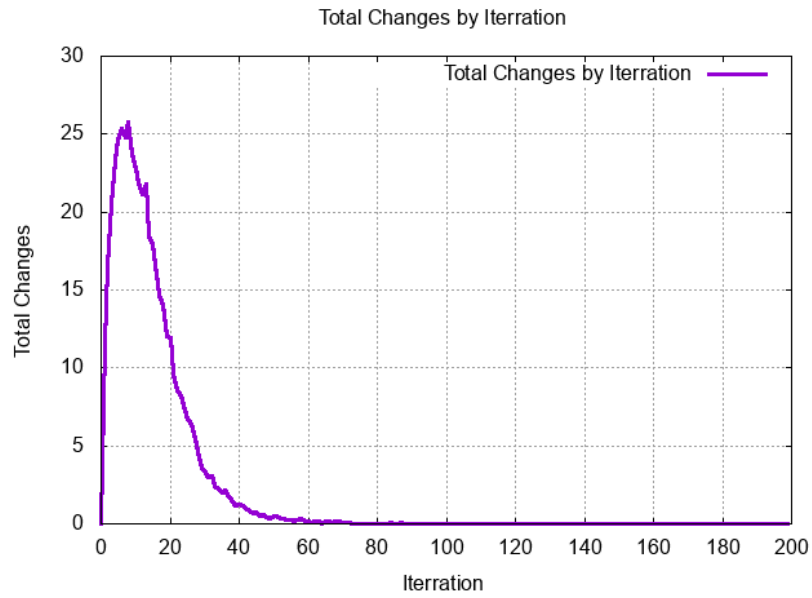


Figure 7: Total Bias Changes by Iteration

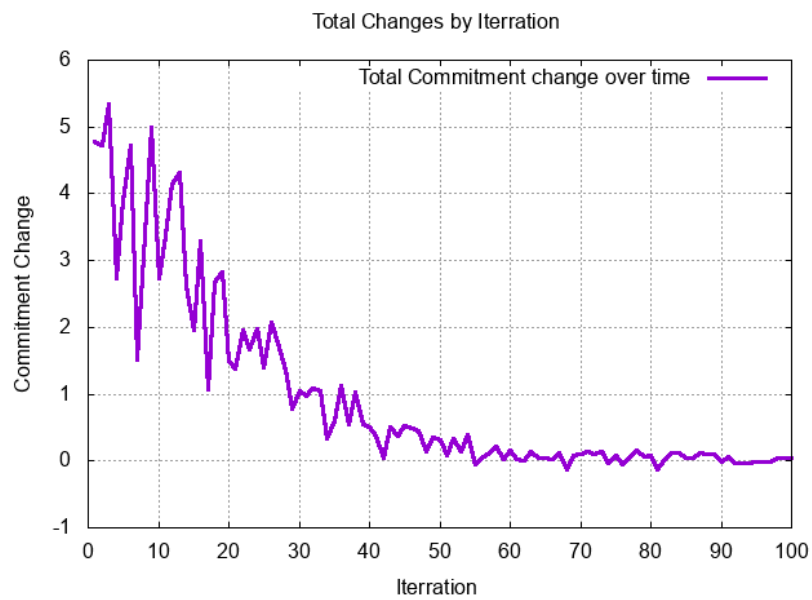


Figure 8: Cumulative Commitment Changes by Iteration

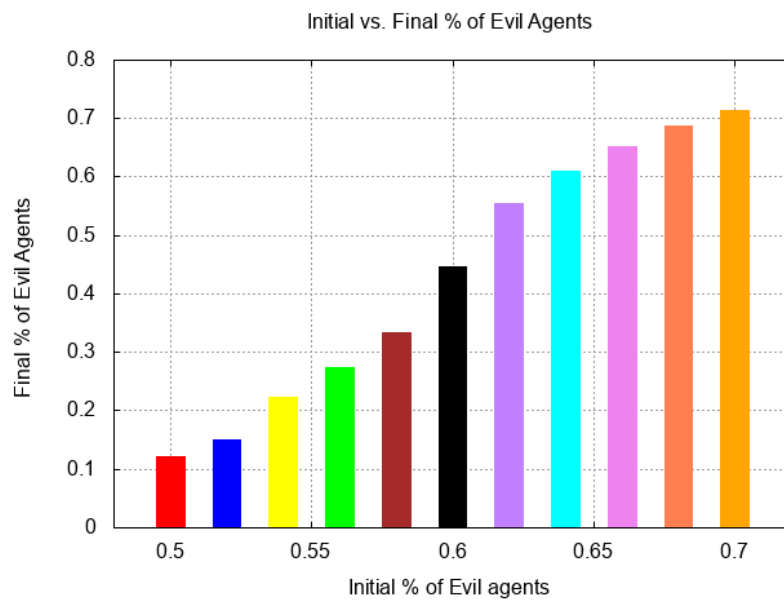


Figure 9: Beginning v. Ending Evil Population Percentage. Generated with varying starting evil percentages. Initial evil commitment levels were generated between 0.0 and 0.5, while initial good commitment levels were generated between 0.5 and 1.0. This gives the good agents a commitment advantage but are still in the minority.