



# Lyra: Simulating Believable Opinionated Virtual Characters

Sasha Azad



Principles of  
Expressive Machines

# OUTLINE

## Motivation

## Evaluation



Motivation

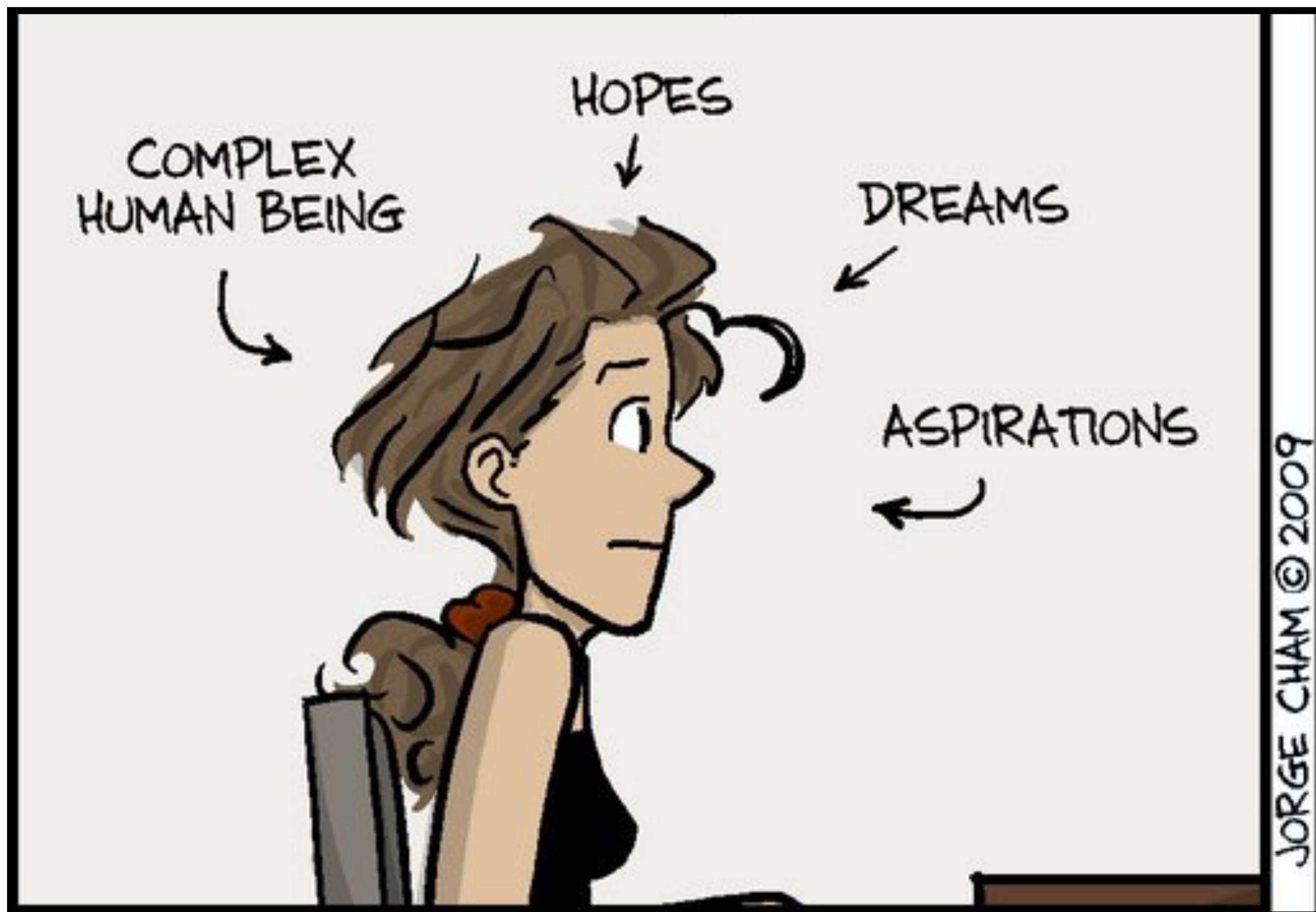
Related Work

System Goals

**Motivation**

Lyra Model and Simulation

Evaluation



Motivation

Related Work

System Goals



## Motivation

Motivation

## Related Work

Lyra Model and Simulation

## System Goals

Evaluation



@PEANUTSSPECIALS

Motivation

Related Work

System Goals

# Opinion Dynamics

- Group formation - social scientists, historians, psychologists etc
- (field) "**Computer Scientists** work to fix easily fooled AI."
- (region) "the **Scottish** voted to overwhelmingly remain in the referendum."

(political ideology) Democrats (US), Tories (UK)  
(fans) Whovians (show), Potterheads (book), Beatlemaniacs (music)

"Individuals relating to a group is an ongoing process of uncertain, fragile, controversial and ever-shifting ties." (Latour 2005)

Motivation

Related Work

System Goals

Motivation

Lyra Model and Simulation

Evaluation

# Opinion Dynamics

- **Scottish, Computer Scientists, Democrats, Whovians**
- Form their own social rules / templates
- Interactions that go against the group's values would be looked upon unfavourably by group members
- Adhere to recognisable social practices and enculturated responses
- Subscribe to sources of information
- Form meaningful connections with group members

Motivation

Related Work

System Goals

# Related Work: Believable NPCs

## Prior Work

- ▶ **Measuring believability**  
Togelius 2013; Thomas 1981; Champandard 2003;  
Bateman 2005
- ▶ **Authoring narratives for various geo-locations**  
Macvean 2011; Dow 2006
- ▶ **Allow NPCs to reason and plan to achieve their goals**  
Leepus 2014; Kunda 1990; Cavazza 2002
- ▶ **Express knowledge and belief**  
Ever 2018; Rowe 2008

## Lyra

- ▶ **Accounting for regional, cultural biases**
- ▶ **Accounting for reasoning under partisanship**
- ▶ **Produce dialog modifiers that indicate the opinions and belief**

Motivation

Related Work

System Goals

Motivation

Lyra Model and Simulation

Evaluation

# Related Work: Social Simulation

## Prior Work

### ▶ Measuring believability

Afonso 2008; Swartout 2006; Riedl 2016;  
Warpefelt 2016

### ▶ Social Practices Templates

Mosher 2006; Mateas 2005; Evans 2013; Wang 2007

### ▶ Social Physics Architecture Model

McCoy 2010; Latour 2005

### ▶ Dynamic Opinion Modeling

Wang 2014; Asch 1955;

## Lyra

### ▶ Computational Social Simulation + Narrative Intelligence

### ▶ Social practices and rules emerge

### ▶ Social relationships affected by opinions held

Motivation

Related Work

System Goals

Motivation

Lyra Model and Simulation

Evaluation

# Related Work: Measuring Believability

- ▶ Game believability is a critical subcomponent of player experience (Togelius 2013)
  - ▶ Linked to stream of player emotions triggered by events during interaction
  - ▶ Linked to cognitive and behavioural processes incited during gameplay
- ▶ Characters whose adventures and misfortunes make people laugh and cry... it's what creates the illusion of life. (Thomas 1981)
- ▶ Appearance of human intelligence or human-likeness adds value to an NPC and to quality of gameplay (Togelius et al. 2013; Champandard 2003; Bateman and Boon 2005)

# Lyra Goals

## System Goals

## Evaluation

Generic Knowledge Model

- Be used for a wide variety of datasets or topics discussed
- Be able to represent the source and an initial rating of the information

Motivation

Related Work

System Goals

Motivation

Lyra Model and Simulation

Evaluation

# Lyra Goals

## System Goals

## Study Goals

Generic Knowledge Model

Accounting for Bias

- Inherent bias in characters on topic
- Bias from the information source
- Allow NPCs to subscribe / unsubscribe to sources of information over time (feed/starve NPC's inherent bias)

Motivation

Related Work

System Goals

Motivation

Lyra Model and Simulation

Evaluation

# Lyra Goals

## System Goals

## Study Goals

Generic Knowledge Model

Accounting for Bias

Discussion Model

- Communicate and influence each other's views
- Ad-hoc groups and relationships forming during social interactions

Motivation

Related Work

System Goals

Motivation

Lyra Model and Simulation

Evaluation

# Lyra Goals

## System Goals

## Study Goals

Generic Knowledge Model

Accounting for Bias

Discussion Model



Motivation

Related Work

System Goals

Motivation

Lyra Model and Simulation

Evaluation



# Addressing the Elephant in the Room: Opinionated Virtual Characters

*Sasha Azad and Chris Martens, AAAI AIIDE Workshop on  
Experimental AI in Games (EXAG), 2018.*

Knowledge   Bias   Simulation

Motivation

Lyra Model and Simulation

Evaluation

# Generic Model of Knowledge

## Rating

- The personal judgment, favour or measure of impartiality associated

Example: Ratings for a show, reviews for a paper, bias for media source

## Topics

- A clustering of information in a specific subject, or field of information.

Example: Sci-Fi, artificial intelligence, gun control

Knowledge

Bias

Simulation

Motivation

Lyra Model and Simulation

Evaluation

# Objects of Discussion

- Single unit of information chosen to debate
- New information: Note the original authorial rating, own views on topic

Example: Doctor Who, procedural content generation, news article

# Sources

- Create information on objects of discussions and topics
- Sources may have a rating, representing the expected rating (or bias) of the information they produce

Example: Rotten Tomatoes, AAAI, New York Times

Knowledge

Bias

Simulation

Motivation

Lyra Model and Simulation

Evaluation

# Discussion Datasets

Topics	Objects of Discussion	Sources	Rating
Political Issues e.g. Immigration	News articles	Online or Print Media	Political Bias or Affiliation
Political Issues e.g. Immigration	Political candidates	Articles, Interviews, Candidate Rally	Approval Ratings
Research Topics e.g. AI, Games	Conference Papers	Journals, Conference Proceedings	Journal or Conference Rankings
Film Genres e.g. Fantasy, Sci-Fi	Movies	Movie Studios	Rotten Tomatoes ratings

Knowledge      Bias      Simulation

# Accounting for Bias

## Attitude

- Agent's **private** views on a specific issue [-1, 1]
- TV Shows: [Hate, Love]; Politics: [Left, Right]; Reviews: [Reject, Accept]

## Opinion

- Agent's **outwardly expressed or shared** views on an issue [-1, 1]
- Can be different from attitude

*Wang (2014); Hegselmann (2002); Asch (1955)*

Knowledge      Bias      Simulation

Motivation

Lyra Model and Simulation

Evaluation

# Bias

- Agent's predisposition to adopt a particular view
- Bias informed by:
  - Own or inherited views
  - Initial bias imparted from the introduction of the topic

# Uncertainty

- A measure of an agent's confidence in their view
- The higher the uncertainty, the more likely the agent is to change their mind or accept other perspectives

Knowledge

Bias

Simulation

Motivation

Lyra Model and Simulation

Evaluation

# Public Compliance Threshold

- Allows agent to feel accepted within the community
- When the strength of the public opinion exceeds this value, the agent will choose to comply with the public opinion

# Private Acceptance Threshold

- Allows agent to stand ground, or stick to their own views
- When the strength of the public opinion is below this value, the agent will stand ground

Knowledge

Bias

Simulation

Motivation

Lyra Model and Simulation

Evaluation

# Lyra Simulation

## Assigning Initial Cultural Bias

- Assign cultural bias across population based on some attribute
- Children inherit as bias the mean of their parent's biases
- May change these attitudes over time through conversations with other dialogists



Knowledge      Bias      Simulation

Motivation

Lyra Model and Simulation

Evaluation

# Discussion Algorithm

- Cluster all expressed opinions from participants (Jenks 1967)
- Check for public consensus
- Check for presence of normative social influence (peer pressure)
- Realign character views for participants



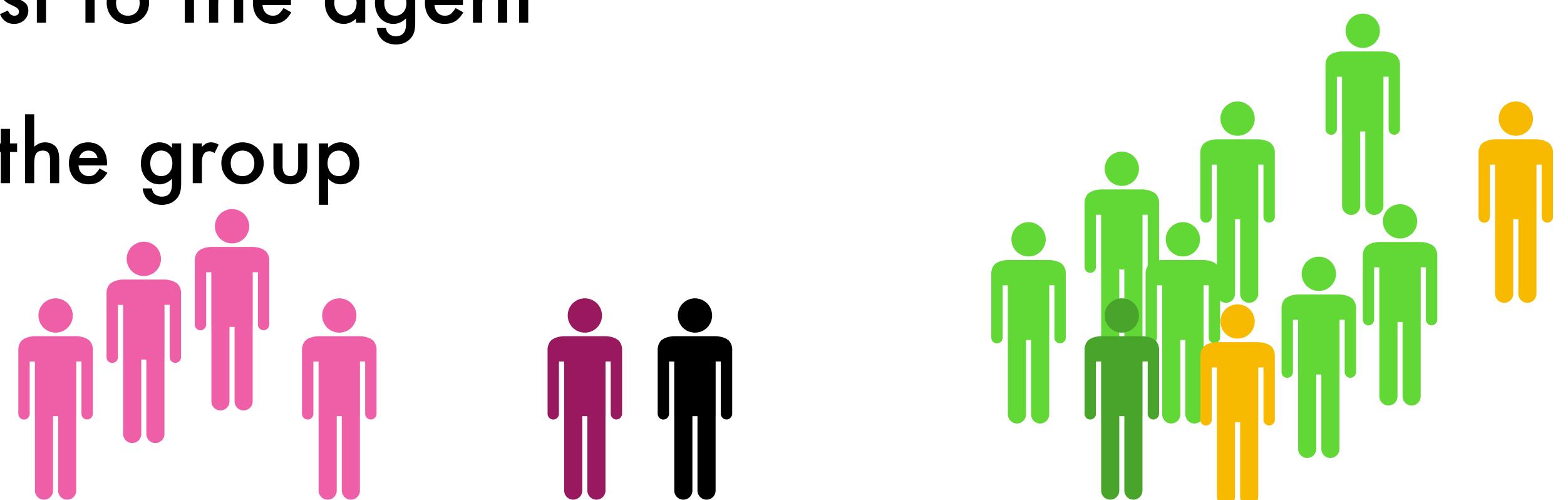
Knowledge

Bias

Simulation

# Public Consensus Formed

- Agents with **high uncertainty**
  - Realign views to that of the largest opinion group
- Agents with **low uncertainty**
  - Find group with opinion closest to the agent
  - Calculate opinion strength of the group



Knowledge

Bias

Simulation

Motivation

Lyra Model and Simulation

Evaluation

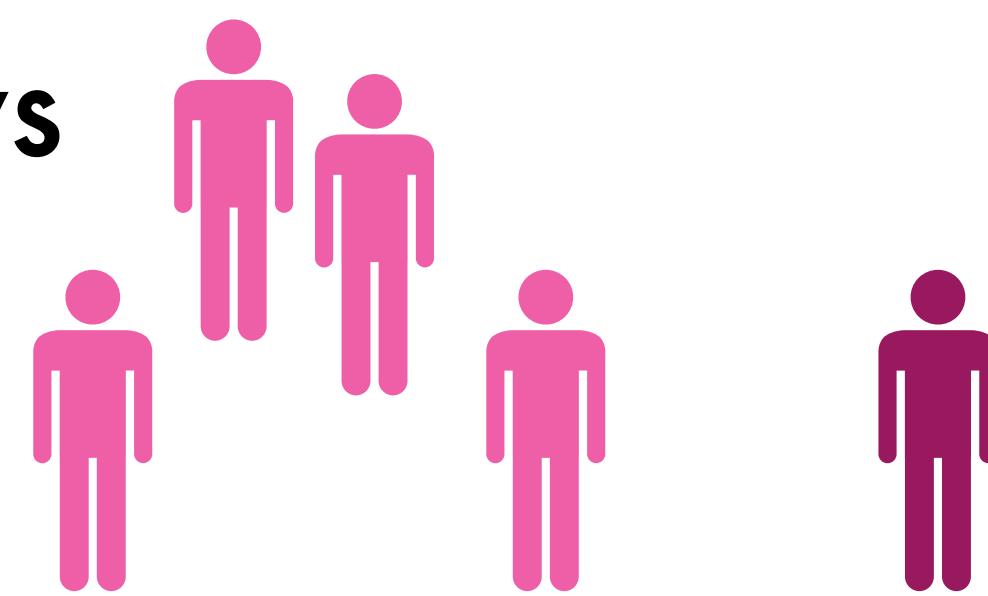
# Opinion Strength

## Group Factors

- Size of the group
- Homogeneity of the opinions in the group (variance)

## Agent Factors

- Discrepancies in the agent's opinions and attitude
- Uncertainty in the agent's own views



Knowledge

Bias

Simulation

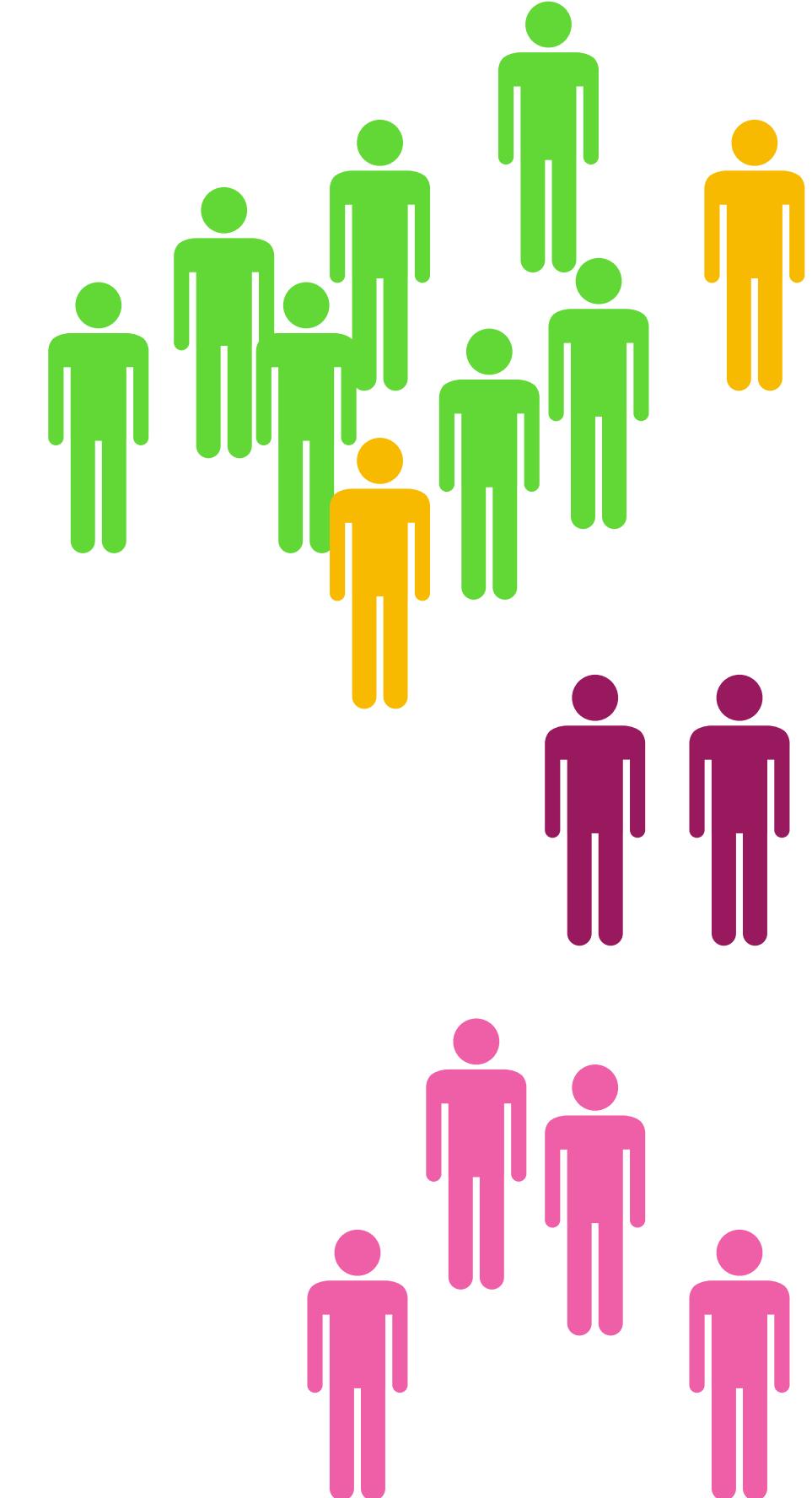
Motivation

Lyra Model and Simulation

Evaluation

# Public Consensus Formed

- **Low op\_str:** The agent does not change their mind
- **Moderate op\_str:**
  - Low uncertainty - Agents believe that the change in their views are a natural and expected evolution
  - High uncertainty - Concede the conversation, realign their views to match.
- **High op\_str:** Recognise peer pressure. Realign opinion, but not attitude. Increase the uncertainty in views.



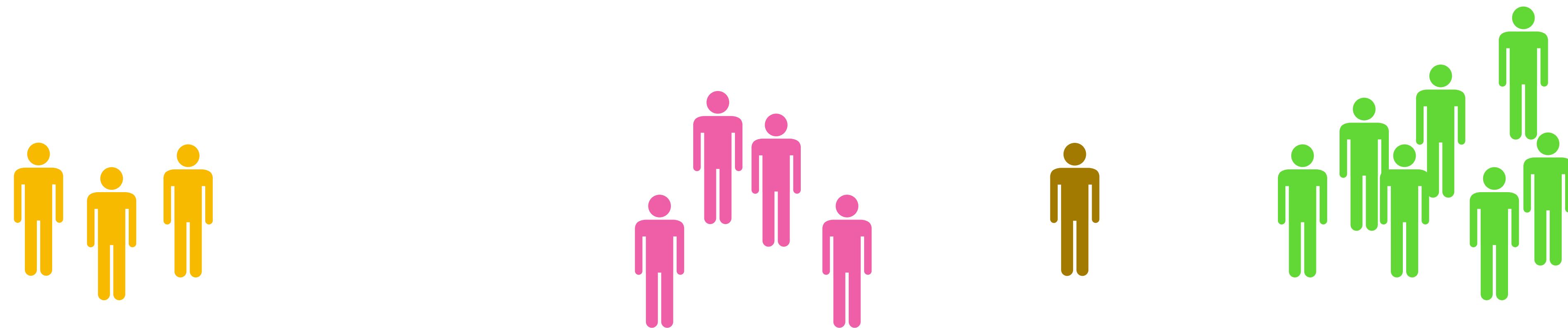
Knowledge

Bias

Simulation

# No Public Consensus Formed

- Find cluster of opinions most similar to that of the NPC
- Realign opinions and attitudes to the mean of the cluster



Knowledge

Bias

Simulation

Motivation

Lyra Model and Simulation

Evaluation

# Realign General Attitudes

- Find new alignment for attitudes and opinions for topics and sources
- Subscribe to new sources and/or unsubscribe from old ones
- Update relationship with group participants



Knowledge

Bias

Simulation

Motivation

Lyra Model and Simulation

Evaluation

# OUTLINE

## System Goals

Generic Knowledge Model

Accounting for Bias

Discussion Model

## Evaluation Goals

**Designing legible simulation output**

**Evaluate conversations with a human subject study**

**Extract insights from study to inform future research**



Output Legibility

Study Design

Analysis

Motivation

Lyra Model and Simulation

Evaluation

# EVALUATION GOALS



## Designing legible simulation output

Generate descriptions to follow an NPC's reasoning

- Choice of domain & scale
- Dealing with authoring bias
- Graphical & Textual descriptors



## Evaluate conversations with a human subject study

Evaluate the generated conversations with a human subject study

- Study Design
- Methods



## Extract insights from study to inform future research

Extract insights from the study on

- Believability & Political bias
- Believability & test conditions
- Clustering evaluation
- Qualitative believability analysis

# Legible Simulation Output

## Choice of Conversational Domain

- Familiar, relatable domain for target demographics
- Quantifiable metric of positions
- Imagine NPC dialogues to sway others to their perspectives
- Should be able to judge clusters and coalitions of like minded NPCs



Output Legibility

Study Design

Analysis

Motivation

Lyra Model and Simulation

Evaluation

# Features of the AllSides Dataset

## Democrats Predict Expedited Impeachment Probe

House Democrats are preparing for a quick impeachment inquiry into President Donald Trump. Recently subpoenaed Secretary of State Mike Pompeo has until Oct. 4 to hand records of Trump's...

### From the Right

Polls shifting in favor of impeaching Trump

Washington Examiner 

### From the Left

Democrats count on Schiff to deliver focused impeachment inquiry of Trump

Washington Post 

### From the Center

Nancy Pelosi has put the Trump impeachment inquiry on a fast track. Here's the plan,...

USA TODAY 

- API accessing corpus
- Clustered by issues
- Tagged with bias

# Example Discussion

- **Object of Discussion:** Discussion on news article “Room for Debate: Should ‘Birthright Citizenship’ Be Abolished”
- **Source:** NY Times (Bias: Leaning Left)
- **Where:** At work with colleagues
- **Topic:** Immigration
- **Duration:** 11 minutes
- **Number of participants:** 4

Output Legibility

Study Design

Analysis

Motivation

Lyra Model and Simulation

Evaluation

# Example Discussion

Discussion for 11 minutes on "Room for Debate: Should 'Birthright Citizenship' Be Abolished?"

Source: NYTimes

Participants: Richard Cain, Ruth Franklin, Vickie Nguyen-Self, Suzanne Sorenson

## Views:

Richard: less uncertain.

Old  $\rightarrow$  att: -0.5 | op: -0.559 | unc: 0.632

New  $\rightarrow$  att: -0.5 | op: -0.559 | unc: 0.554

Vickie: change in views, less uncertain.

Old  $\rightarrow$  att: 1 | op: 0.535 | unc: 0.464

New  $\rightarrow$  att: 0.948 | op: 0.521 | unc: 0.497

Ruth: no change in views.

att: 0.0 | op: -0.063 | unc: 0.225

att: 0.0 | op: -0.063 | unc: 0.225

Suzanne: no change in views

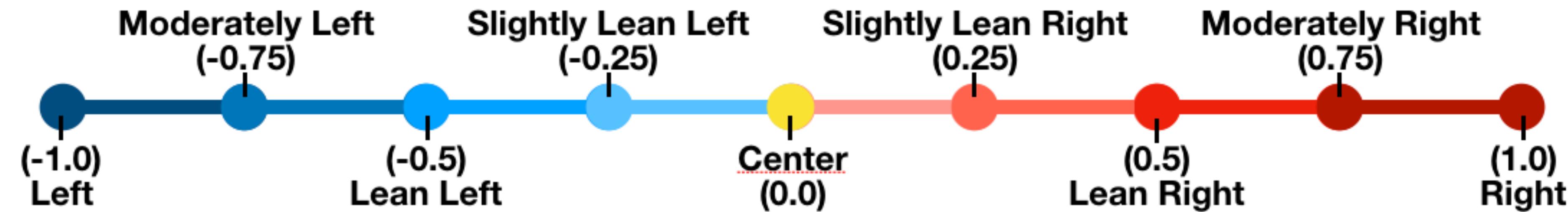
att: -0.5 | op: -0.559 | unc: 0.059

att: -0.5 | op: -0.559 | unc: 0.059

# Following the change in NPC views

- ▶ Hard to relate to the numerical change in character opinions
- ▶ Solution: Simplified Political Scale

Ruth: no change in views.  
 att: 0.0 | op: -0.063 | unc: 0.225



# Authoring Bias for Dialogues

- ▶ Authoring dialogue to go with a character's views untenable
- ▶ Solution: Generate textual descriptions

Too many differing opinion groups present. Public Opinion not formed on the matter.

Ada Lawson did not agree with the other opinions.

They realized their expressed opinions did not truly match their internal attitudes.

They tried to reconcile the difference.

Ada Lawson updated their view rating

Ashley Thurston was swayed by Helga Bass's argument.

They decided to change their rating to indicate the same.

Ashley Thurston updated their view rating

# Authoring Bias for Dialogues

- ▶ Authoring dialogue to go with a character's views during a round untenable
- ▶ Solution: Generate textual descriptions

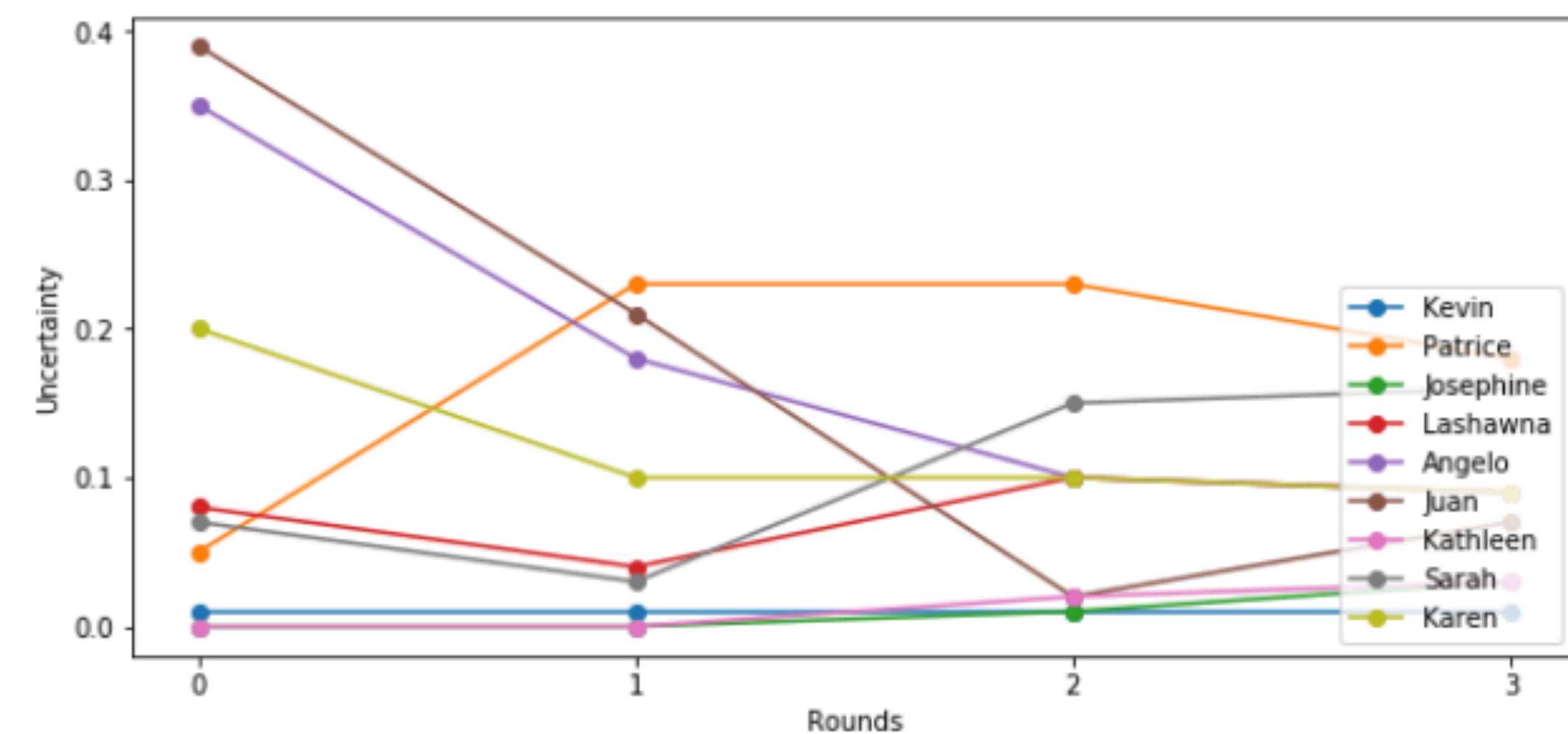
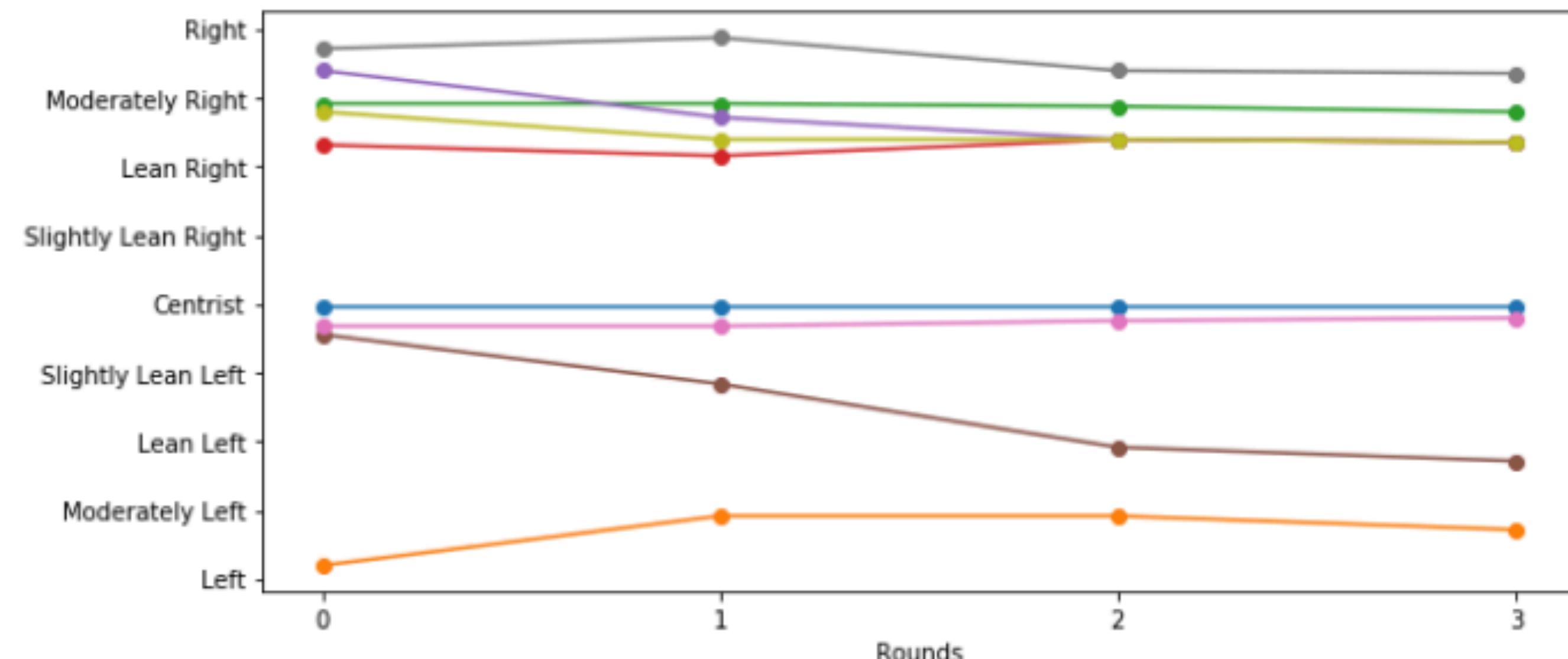
Ada Lawson realized the opinion they expressed was inconsistent with their internal attitude on the article. They looked for the group with views closest to their own expressed opinions. The closest group was the one with Johnnie Helm.

Ada Lawson thought about whether the group opinion was strong enough. After an internal debate Ada Lawson realized that the strength of the group's convictions was too weak.

Ada Lawson did not change their mind.

# Graphical Descriptions

- ▶ Descriptions lengthy,
- ▶ Too many variables to track
- ▶ Solution: Generate chart based descriptions to accompany text



# EVALUATION GOALS



## Designing legible simulation output

Generate descriptions to follow an NPC's reasoning

- Choice of domain & scale
- Dealing with authoring bias
- Graphical & Textual descriptors



## Evaluate conversations with a human subject study

Evaluate the generated conversations with a human subject study

- Study Design
- Methods



## Extract insights from study to inform future research

Extract insights from the study on

- Believability & Political bias
- Believability & test conditions
- Clustering evaluation
- Qualitative believability analysis

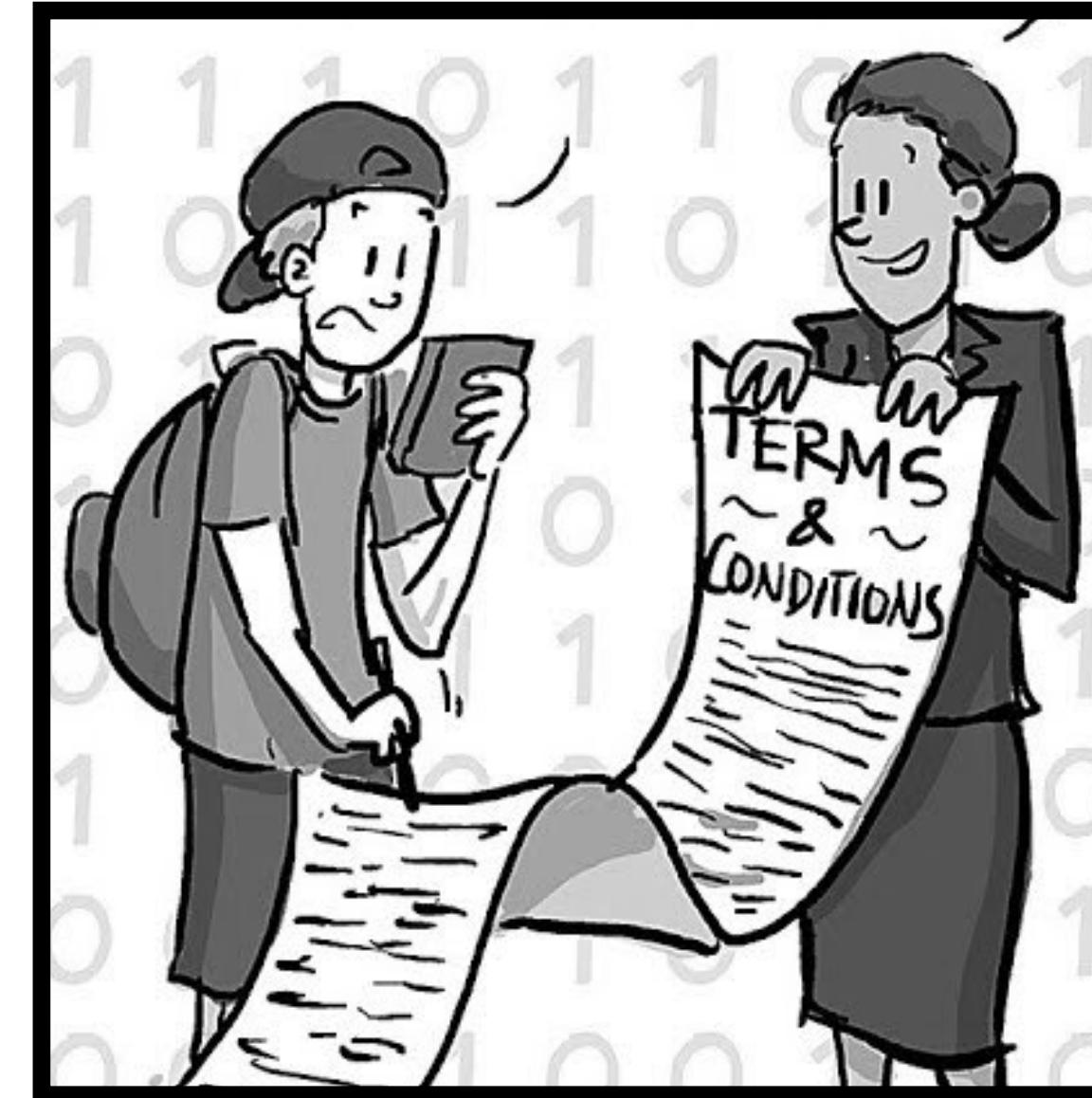
# EVALUATION GOALS



## Designing legible simulation output

Generate descriptions to follow an NPC's reasoning

- Choice of domain & scale
- Dealing with authoring bias
- Graphical & Textual descriptors



## Evaluate conversations with a human subject study

- Evaluate the generated conversations with a human subject study**
- **Study Design**
- **Methods**



## Extract insights from study to inform future research

- Extract insights from the study on
  - Believability & Political bias
  - Believability & test conditions
  - Clustering evaluation
  - Qualitative believability analysis

# Response Demographics

- 21 respondents
- 11 male | 8 female | 1 Other
- 4 PhD | 11 Masters | 4 Bachelors | 1 Associate | 1 College credit
- 16 Liberal | 4 Conservative | 1 Declined to reply
- Views on immigration and gun control

Output Legibility

Study Design

Analysis

Motivation

Lyra Model and Simulation

Evaluation

# Study Design

- Discussion parameters: group size, conversation duration
- Queries (per discussion): 1 (*Not Believable At All*) – 5 (*Very believable*)
  - Believability Rating
    - Most Believable
    - Least Believable
    - Reasoning Queries
- Clustering Analysis

*Open Coding /  
Qualitative Reasoning*

Output Legibility

Study Design

Analysis

Motivation

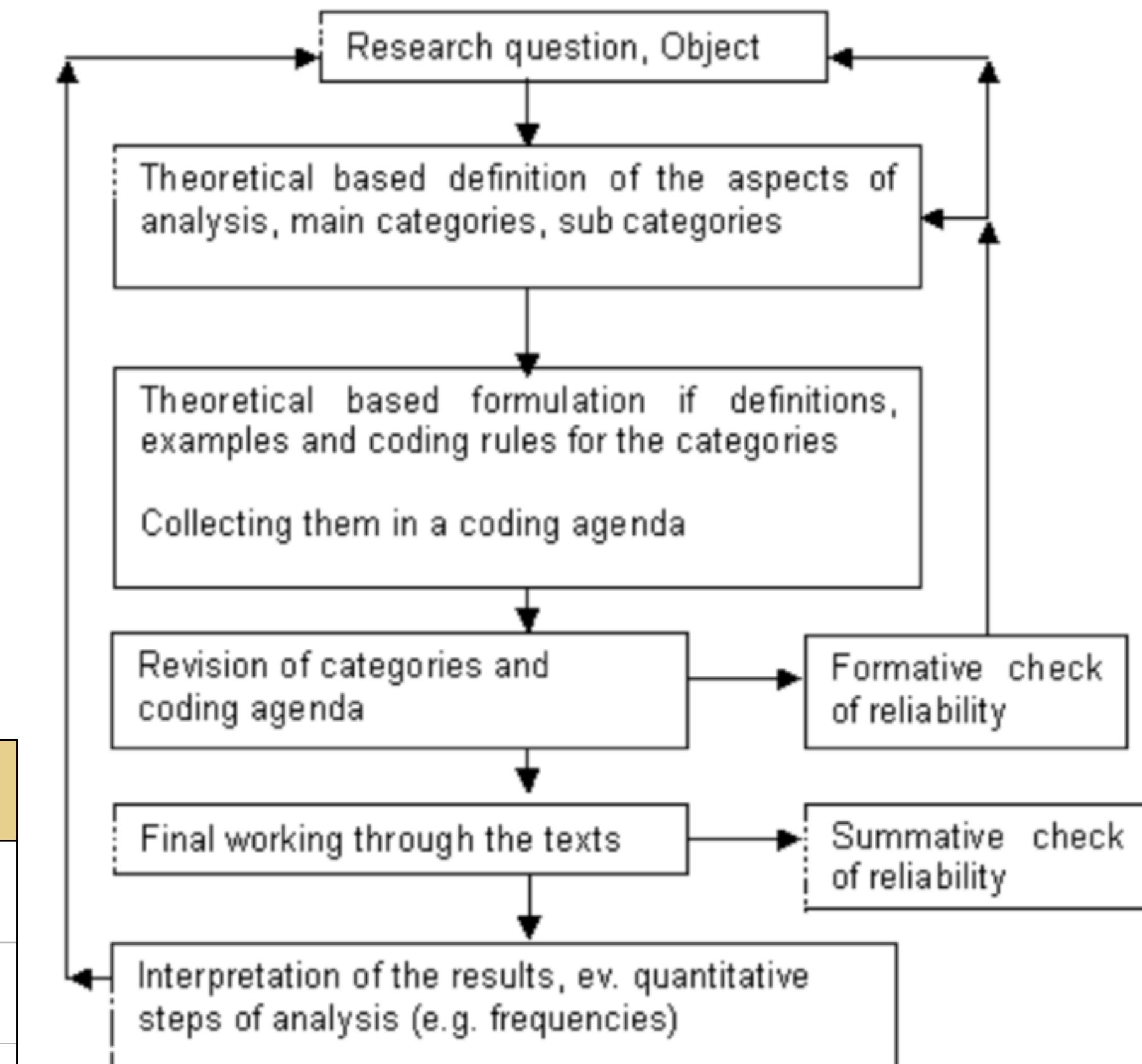
Lyra Model and Simulation

Evaluation

# Coding Scheme - Creation and Validation

- Directed Content Analysis
- Open / Thematic Coding
- Validation of initial coding scheme
- Almost Perfect Agreement
- 44 codes

Measure	Agreement
Fleiss Kappa	0.9099
Cohen Kappa	0.9121
alpha	0.9012



Deductive Category Application (Mayring 2004)

Output Legibility

Study Design

Analysis

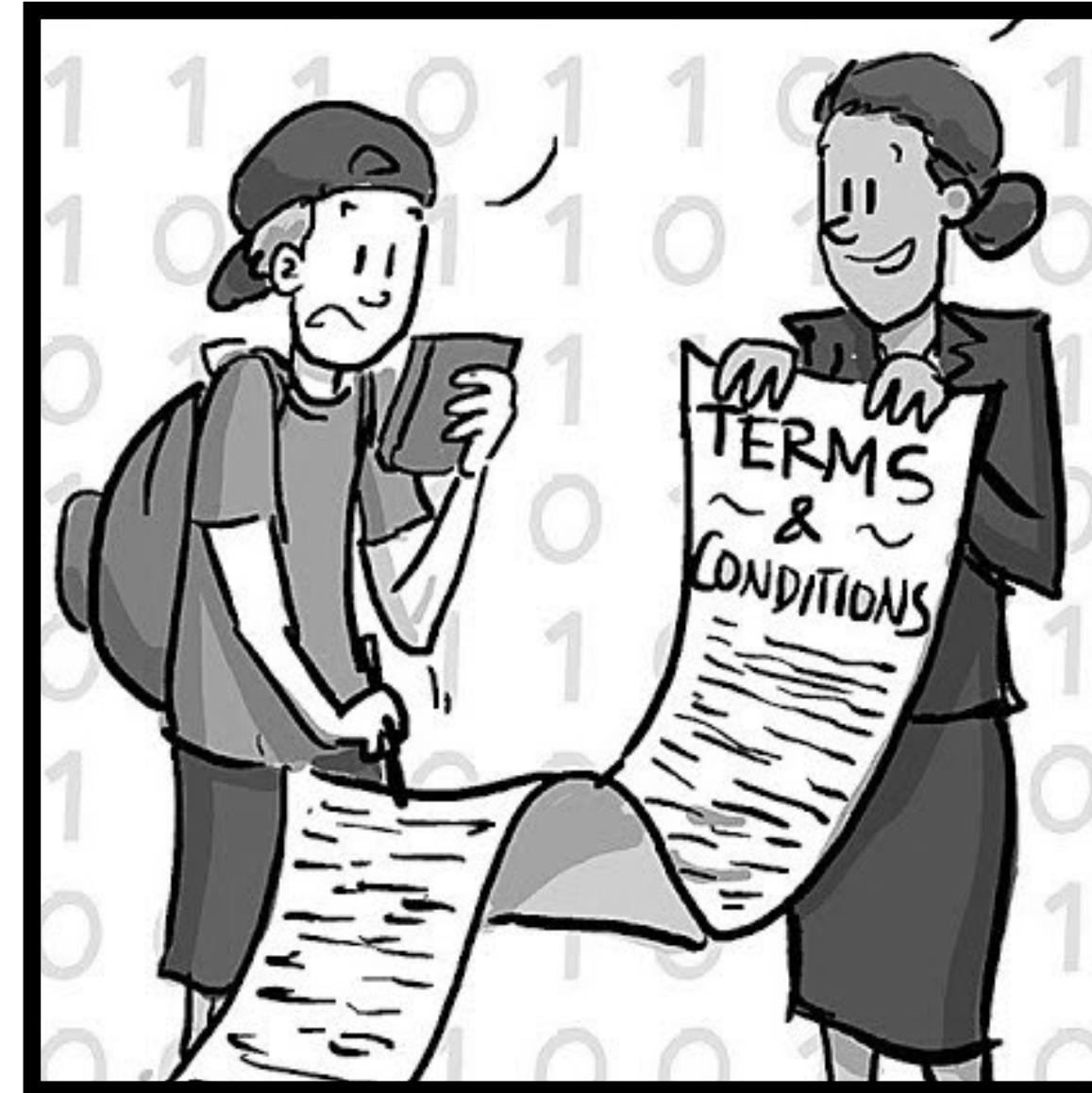
# EVALUATION GOALS



## Designing legible simulation output

Generate descriptions to follow an NPC's reasoning

- Choice of domain & scale
- Dealing with authoring bias
- Graphical & Textual descriptors



## Evaluate conversations with a human subject study

- Evaluate the generated conversations with a human subject study
- Study Design
- Methods



## Extract insights from study to inform future research

Extract insights from the study on

- Believability & Political bias
- Believability & test conditions
- Clustering evaluation
- Qualitative believability analysis

# EVALUATION GOALS



## Designing legible simulation output

Generate descriptions to follow an NPC's reasoning

- Choice of domain & scale
- Dealing with authoring bias
- Graphical & Textual descriptors



## Evaluate conversations with a human subject study

Evaluate the generated conversations with a human subject study

- Study Design
- Methods



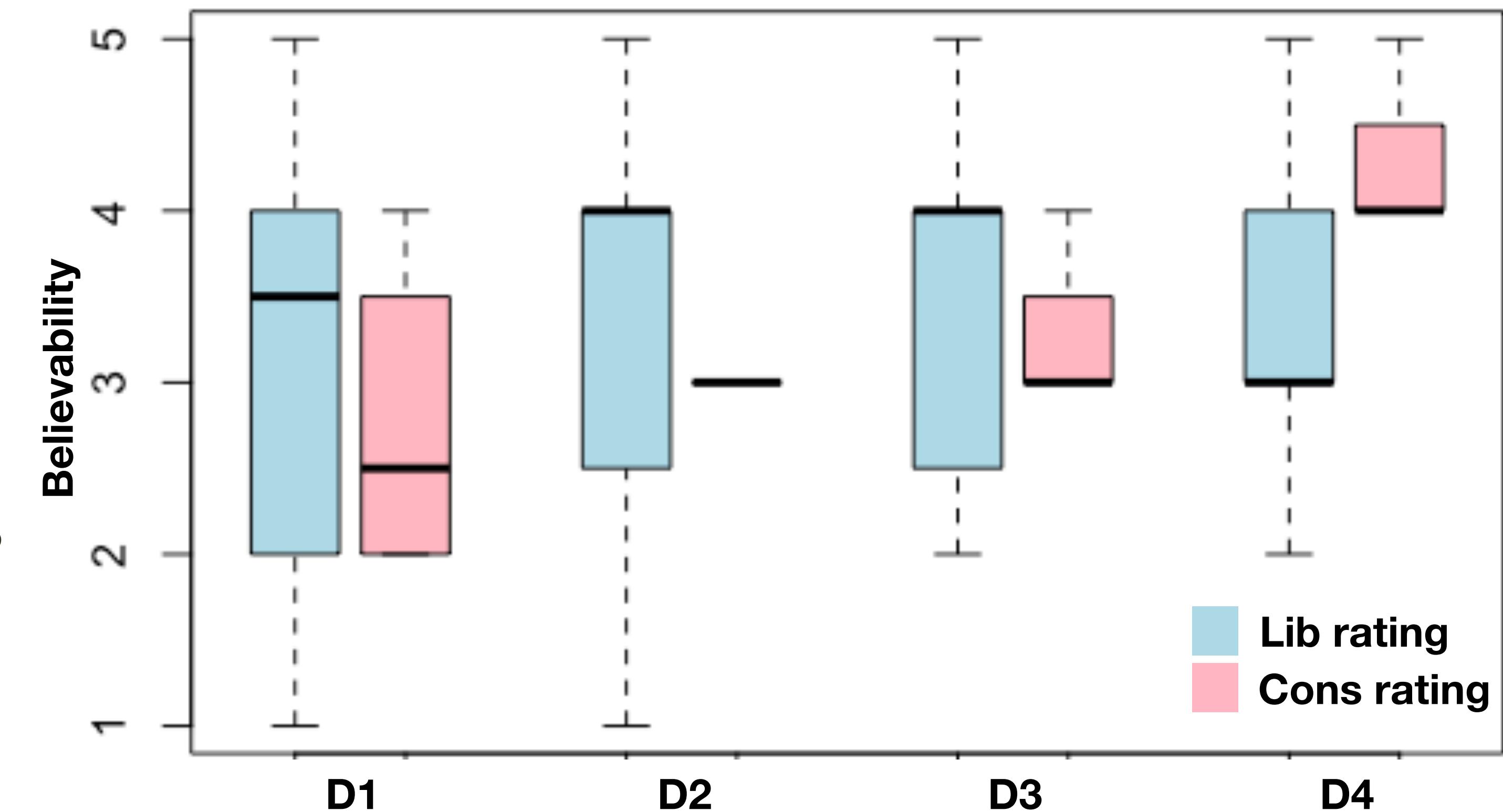
## Extract insights from study to inform future research

Extract insights from the study on

- Believability & Political bias
- Believability & test conditions
- Clustering evaluation
- Qualitative believability analysis

# RQ1: Does the measure of the believability depend on the personal political biases of the respondents?

- Non-parametric Mann-Whitney U test using overall political bias
- Linear regression model using rating and political descriptors
- No significant differences in how Liberals and Conservatives rate discussions



Output Legibility

Study Design

**Analysis**

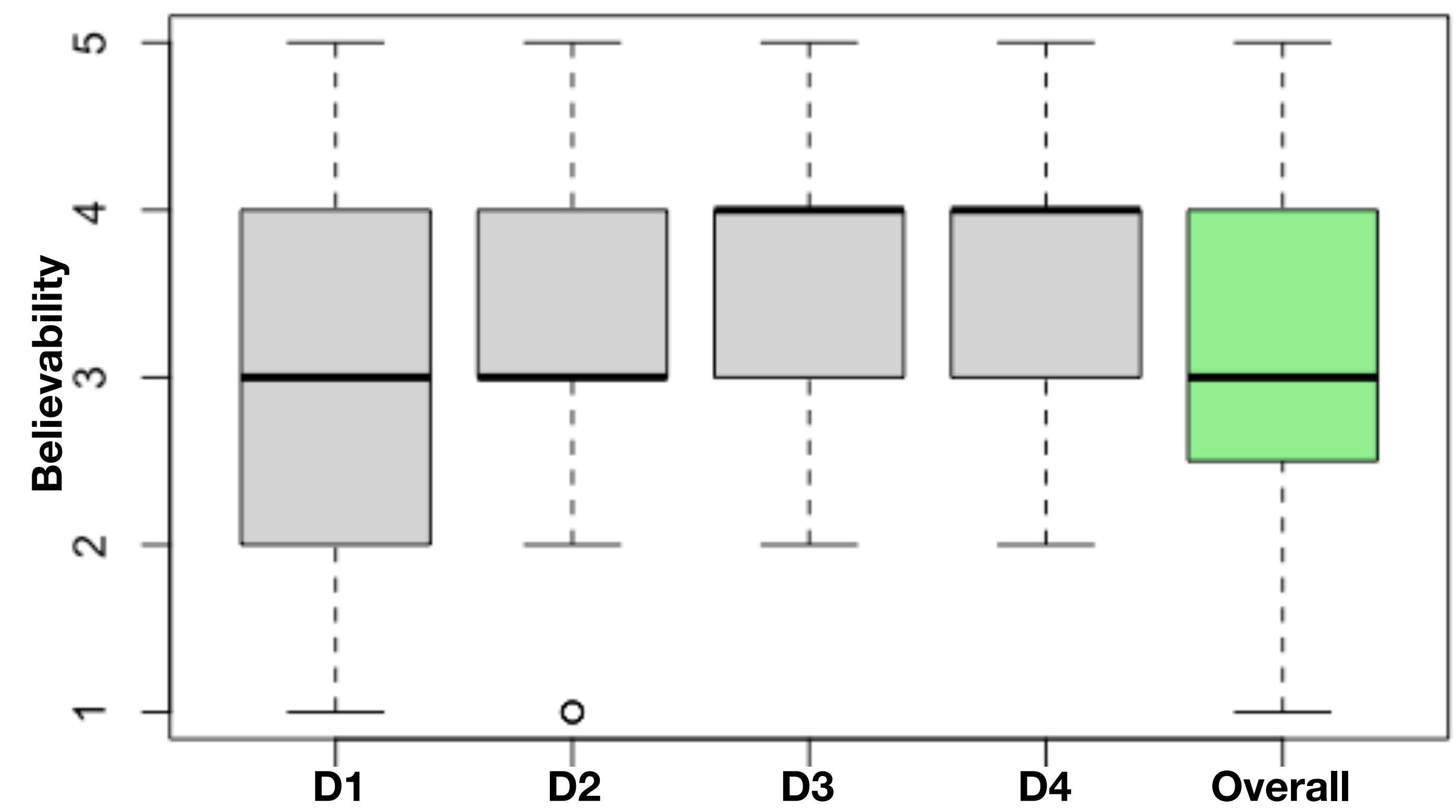
Motivation

Lyra Model and Simulation

Evaluation

## RQ2: Does the measure of believability in the generated conversations vary across conversation parameters?

- Friedman Test: non-parametric alternative to one-way ANOVA with repeated measures
- No significant differences in how conversations were rated across different discussion parameters.



Output Legibility

Study Design

**Analysis**

Motivation

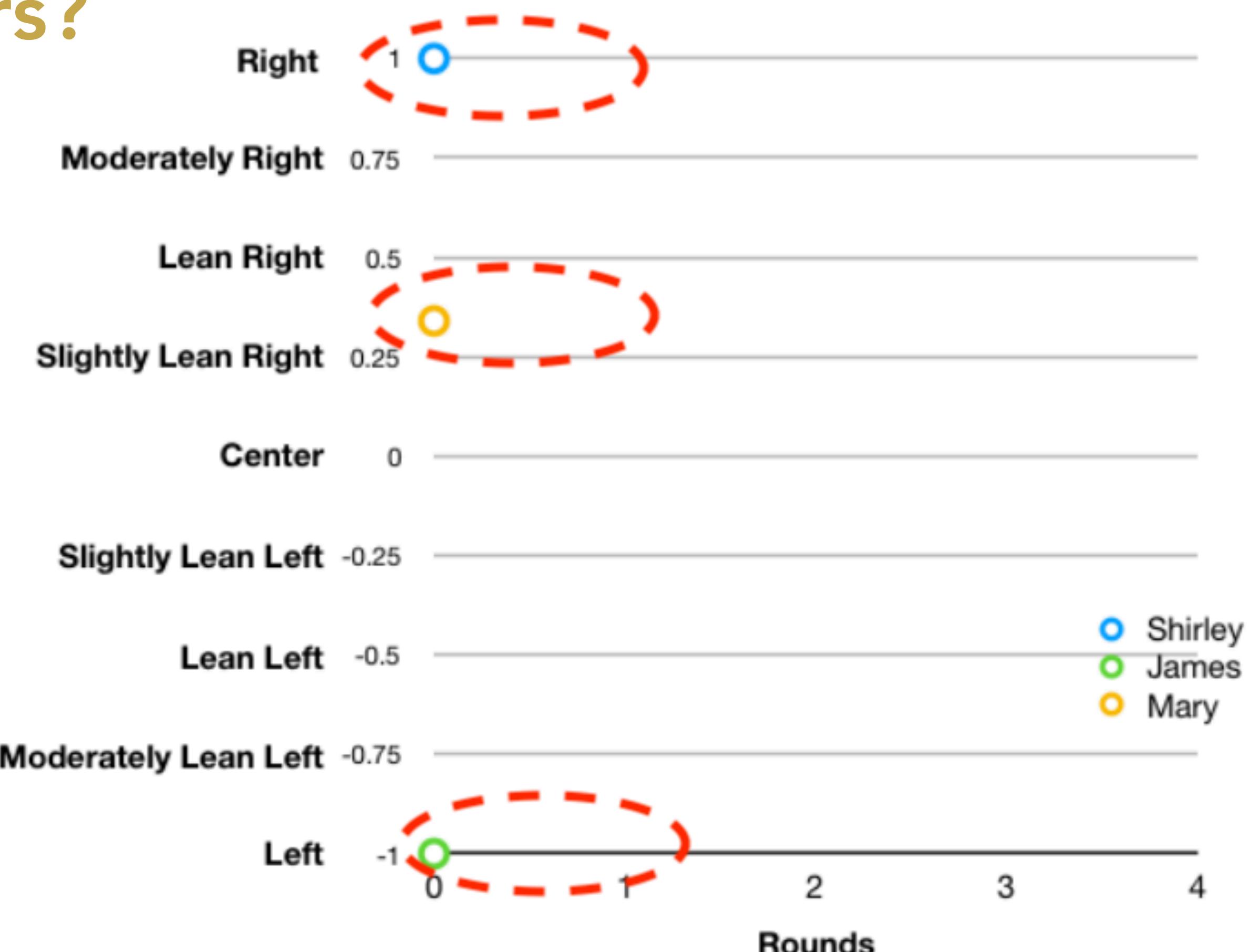
Lyra Model and Simulation

Evaluation

## RQ3: How similar is Lyra's clustering to how humans define and group like-minded virtual characters?

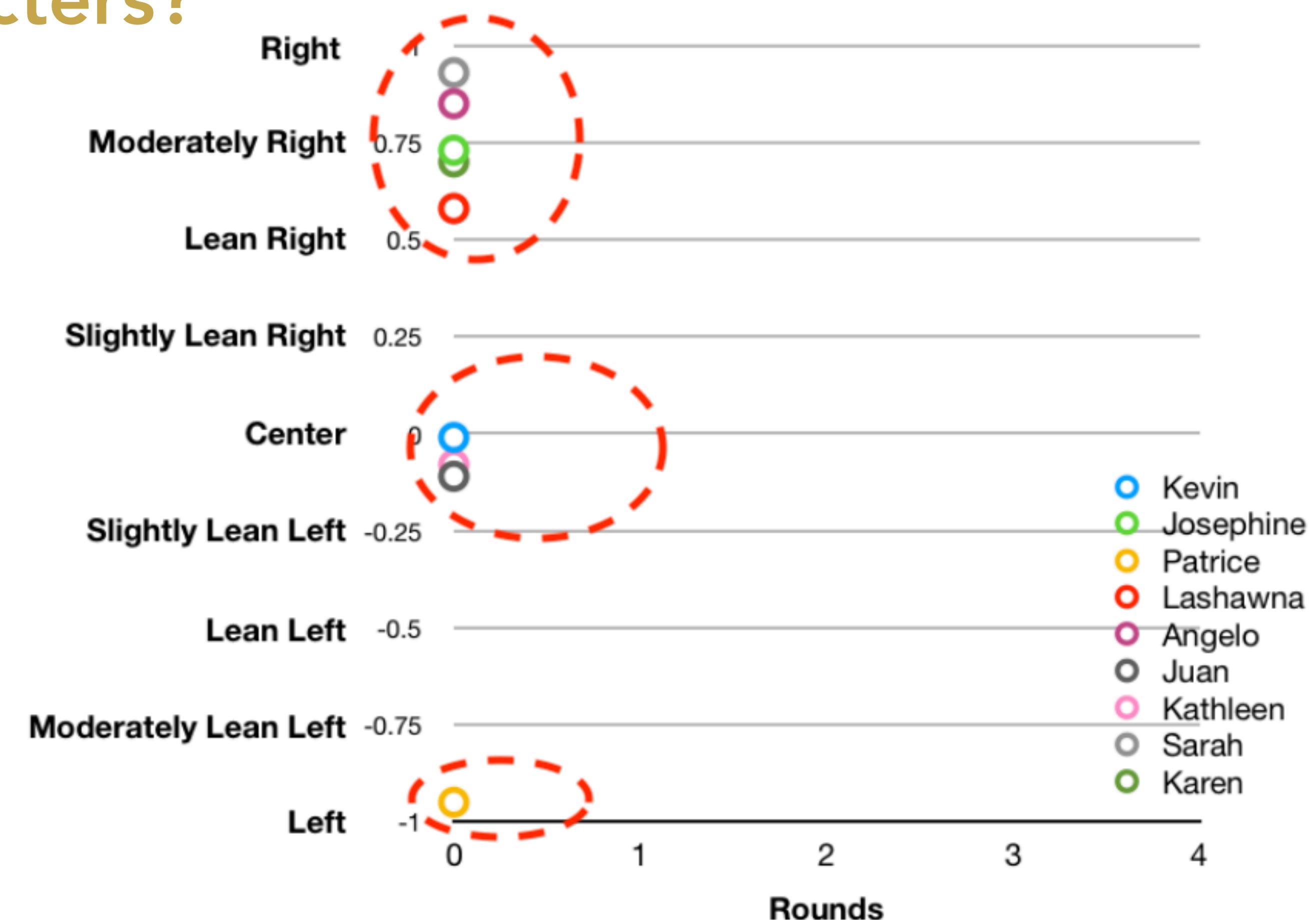
- Jenks Natural Breaks  
 $GVF \geq 0.9$

	<b>Model Agreement</b>	<b>Respondent Agreement</b>
D1	0.1428	0.666
D2	0.5714	0.5714
D3	0	0.238
D4	0	0.333



## RQ3: How similar is Lyra's clustering to how humans define and group like-minded virtual characters?

- Jenks Natural Breaks with  $GVF \geq 0.9$
- D3 (lowest agreement)
  - Round 1: 7 clusters (0%)
  - Round 2: 3 clusters (23%)



Output Legibility

Study Design

Analysis

## RQ4: Does using Lyra impact the believability of the virtual characters?

- Believability

**Moderately believable 3.3/5**

- What was the most believable part of the conversation?
- What was the least believable part of the conversation?
- Reasoning questions:
  - Why do you think Ashley was so uncertain of their views?
  - Why do you think James's uncertainty increased?
  - What does Juan's change in opinion tell you of their private attitude?
  - Why do you think Amy's uncertainty increased after Round 2?

**Open Coding /  
Qualitative Reasoning**

Output Legibility

Study Design

Analysis

Motivation

Lyra Model and Simulation

Evaluation

## RQ4: Does using Lyra impact the believability of the virtual characters?

- What was the **most believable** part of the conversation?
- NPC Mentioned Unprompted
  - Standing Ground: "*Helga started at Left; moved to centrist and then closed at left.*" [D1]
  - Polarization: "*That over time and rounds of arguments consensus develops around two poles of thought; even though within the poles there's a range of opinion/degree of certainty*" [D1]
  - Individual Influence: "*Amy was swayed by Ada*" [D1, D4]

Theme	Frequency
NPC mentioned unprompted	23
Standing Ground	18
Similar views converging	12
Influence from groups	10
Used political affiliation stereotype	9
Influence by an individual	8
Polarization	8

# RQ4: Does using Lyra impact the believability of the virtual characters?

- What was the **most believable** part of the conversation?

- NPC Mentioned Unprompted

- Group Influence [D2, D3]:

- "Lashawna swaying slightly more conservative because she had a very convincing and large group and this would easily move her to similar opinion"
- "The fact that James had not changed drastically on his political opinion but has opened up his opinion to uncertainty seems believable since he is outnumbered in the group."

Theme	Frequency
NPC mentioned unprompted	23
Standing Ground	18
Similar views converging	12
Influence from groups	10
Used political affiliation stereotype	9
Influence by an individual	8
Polarization	8

## RQ4: Does using Lyra impact the believability of the virtual characters?

- What was the **most believable** part of the conversation?

- Similar Views Converge [D1, D2]: "No drastic changes in views but groups did come closer to same opinion on both sides."
- Used Political Affiliation Stereotype
  - "The consistency with which the Right Opinionated people stuck to their stand"
  - "That the centrist didn't change their opinion much"
  - "That the most liberal person would be the person most open to changing their mind"

Theme	Frequency
NPC mentioned unprompted	23
Standing Ground	18
Similar views converging	12
Influence from groups	10
Used political affiliation stereotype	9
Influence by an individual	8
Polarization	8

## RQ4: Does using Lyra impact the believability of the virtual characters?

- ▶ What was the **least believable** part of the conversation?
- ▶ Believable: 6 respondents, "I find it believable"
- ▶ NPC Mentioned Unprompted
  - ▶ Influenced by Article [D1, D2]: "That James (someone who was extreme left) was swayed by the [Centrist] Article."
  - ▶ Standing Ground [D2]: "Shirley was not influenced by the other two in any way"

Theme	Frequency
NPC mentioned unprompted	44
Changed Opinion	19
Decreasing Certainty	11
<b>Standing Ground</b>	<b>10</b>
Believable	6
Influenced by Article	6

## RQ4: Does using Lyra impact the believability of the virtual characters?

- ▶ What was the **least believable** part of the conversation?
- ▶ NPC Mentioned Unprompted
- ▶ Changed Opinion [D3]: "The unexpected move of Juan towards the Left and Patrice's position feels like the kind of strange turn that might happen in a real conversation - in a large enough conversation you will see some people's opinion change"
- ▶ Changed Opinion [D4]: "Kennet wasn't persuaded much at all; shifting to the right seemed weird"
- ▶ Decreasing Certainty

Theme	Frequency
NPC mentioned unprompted	44
Changed Opinion	19
Decreasing Certainty	11
<b>Standing Ground</b>	<b>10</b>
Believable	6
Influenced by Article	6

# RQ4: Does using Lyra impact the believability of the virtual characters?

- ▶ Reasoning Queries
- ▶ Individual Influence [D1, D4]: "She was uncertain to begin with and her group mate; who was the most knowledgeable (ie if no of prior articles read is an indicator of knowledge); was also wavering her convictions"
- ▶ NPC mentioned unprompted: "William was persuasive and swayed Amy"
- ▶ Opinion Attitude Difference [D3]: "He didn't want to seem biased externally so wanted to be portrayed as a centrist; but was privately left-leaning"

Theme	Frequency
Individual Influence	19
NPC mentioned unprompted	15
Opinion Attitude Difference	12
Infer facts not provided	11
Group Influence	10
Certainty Convinces	10
Lacking Support	8
Emotions attributed	7

Output Legibility

Study Design

Analysis

# RQ4: Does using Lyra impact the believability of the virtual characters?

- ▶ Reasoning Queries
- ▶ Smaller Discussion Groups [D1, D2]
  - ▶ Certainty Convinces: "You must assume this is because of Johnnie's certainty" or "The opposition members confidence and articulation was strong"
  - ▶ Lacking Support: "Because of the feeling of being marginalised" or "lack of support from like-minded people"

Theme	Frequency
Individual Influence	19
NPC mentioned unprompted	15
Opinion Attitude Difference	12
Infer facts not provided	11
Group Influence	10
Certainty Convinces	10
Lacking Support	8
Emotions attributed	7

# RQ4: Does using Lyra impact the believability of the virtual characters?

- ▶ Reasoning Queries

- ▶ Longer Discussions [D2, D4]

- ▶ Group Influence: "The opposition had convincing arguments or [that there was a] tendency to want to agree with the majority" or "Temporary bias because of peer-pressure in a group of majority conflicting opinions"

- ▶ Shorter Discussions [D1, D3]

- ▶ Infer Facts: "They support innovation and reform strongly" or "Seem to value the Rights and Interests of the others"

Theme	Frequency
Individual Influence	19
NPC mentioned unprompted	15
Opinion Attitude Difference	12
Infer facts not provided	11
Group Influence	10
Certainty Convinces	10
Lacking Support	8
Emotions attributed	7

# RQ4: Does using Lyra impact the believability of the virtual characters?

- ▶ Reasoning Queries
- ▶ Shorter Discussions [D1, D3]
- ▶ Emotions Attributed:
  - ▶ "Changing one's political identity on an issue isn't an easy task and can result in much internal conflict and therefore high uncertainty"
  - ▶ "Because of the feeling of being marginalised"
  - ▶ "Their competitiveness seemed to be declining"
  - ▶ "Seems to care about the well-being of the others"

Theme	Frequency
Individual Influence	19
NPC mentioned unprompted	15
Opinion Attitude Difference	12
Infer facts not provided	11
Group Influence	10
Certainty Convinces	10
Lacking Support	8
Emotions attributed	7

Output Legibility

Study Design

Analysis

Motivation

Lyra Model and Simulation

Evaluation

# Modeling Social Influence and Simulation

- ▶ **Overton Window:** "Everyone else expressed a more rightward view; making Ashley's view appear more extreme left than it actually was."
- ▶ **Polarization:** "No substantial agreement was reached; which is what you might expect from an argument where people's views start out very highly separated from each other"
- ▶ **Peer Pressure:** Respondents pointed out when NPCs seemed "outnumbered" or "In the minority so probably felt uncertain"
- ▶ **Persuasion:** "deliberation within a group is important, with the right convincing you can change someone's mind" or "there is some power in group mentality"

Output Legibility

Study Design

Analysis

# EVALUATION GOALS



## Designing legible simulation output

Generate descriptions to follow an NPC's reasoning

- Choice of domain & scale
- Dealing with authoring bias
- Graphical & Textual descriptors

- ▶ 17 out of 21 respondents were able to interpret the conversations and use them to reason about NPC behaviour
- ▶ 4 had difficulty following the descriptions provided. "Difficult to align with [my] own mental model of the dynamic. **The graphs help;** but the textual description is pretty poor [and] too abstract."

- ▶ Can produce explainable behaviour that matches the expectations of the reader, allowing them to reason about the conversations
  - Study Design
  - Methods
  - Clustering evaluation
  - Qualitative believability analysis

# EVALUATION GOALS



**Evaluate conversations with  
a human subject study**

Evaluate the generated  
conversations with a human  
subject study

- Study Design
- Methods

- ▶ Described the study design and analysis method
- ▶ Only 21 responses
  - 16 Liberal | 4 Conservative | 1 Declined to reply
- ▶ Data not normally distributed
- ▶ Unable to determine statistical significance
- ▶ Mean believability rating: 3.3 Moderately believable
  - Believability & Political bias
  - Believability & test conditions
  - Clustering evaluation
  - Qualitative believability analysis

# EVALUATION GOALS

---



**Extract insights from study  
to inform future research**

Extract insights from the study on

- Believability & Political bias
- Believability & test conditions
- Clustering evaluation
- Qualitative believability analysis

- ▶ Most respondents expected and interpreted opinion change in the way our algorithm performed it
- ▶ Displayed emotional responses to the conversations:  
"I found it believable but depressing that none [of the NPCs] ultimately changed their minds [on Immigration] at the end of Round 3"
- ▶ Attributed emotions to NPCs of competitiveness, charm, support for reform, care for well-being of population
- ▶ Attributed intentions to NPCs of being open minded, liberal

# IN CONCLUSION

## System Goals

Generic Knowledge Model

Accounting for Bias

Discussion Model

## Evaluation Goals

Designing legible simulation output

Evaluate conversations with a human subject study

Extract insights from study to inform future research



# Believability & Lyra

- ▶ Game believability is a critical subcomponent of player experience (Togelius 2013)
  - ▶ Linked to stream of player emotions triggered by events during interaction
  - ▶ Linked to cognitive and behavioural processes incited during gameplay
  - ▶ Systems with believable elements can elicit emotions in the player
- ▶ Characters whose adventures and misfortunes make people laugh and cry... it's what creates the illusion of life.
- ▶ Appearance of human intelligence or human-likeness adds value to an NPC and to quality of gameplay (Togelius et al. 2013; Champadard 2003; Bateman and Boon 2005)

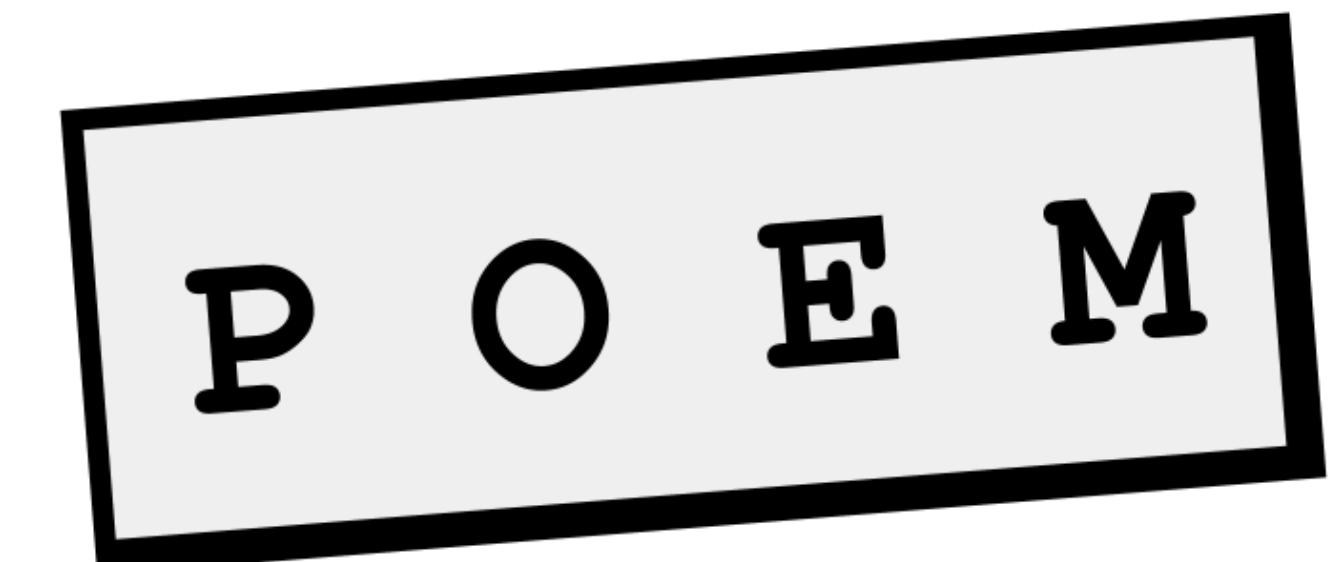
# Believability & Lyra

- ▶ Game believability is a critical subcomponent of player experience (Togelius 2013)
  - ▶ Linked to stream of **player emotions** triggered by events during interaction
  - ▶ Linked to **cognitive and behavioural processes** incited during gameplay
  - ▶ Systems with believable elements can **elicit emotions** in the player
- ▶ Characters whose adventures and misfortunes make people laugh and cry... it's what creates the illusion of life.
- ▶ Appearance of **human intelligence or human-likeness** adds value to an NPC and to quality of gameplay (Togelius et al. 2013; Champadard 2003; Bateman and Boon 2005)



# Lyra: Simulating Believable Opinionated Virtual Characters

Sasha Azad



Principles of  
Expressive Machines

## Most Believable Quotes

## Encoded

1-9: That over time and rounds of arguments consensus develops around two poles of thought; even though within the poles there's a range of opinion/degree of certainty.

#Polarization #SimilarViewsConverge  
#Believable #Expected  
#IdentifyingSimilarGroups  
#ClusteringBelievable

1-12: Right leaning viewpoints stayed right

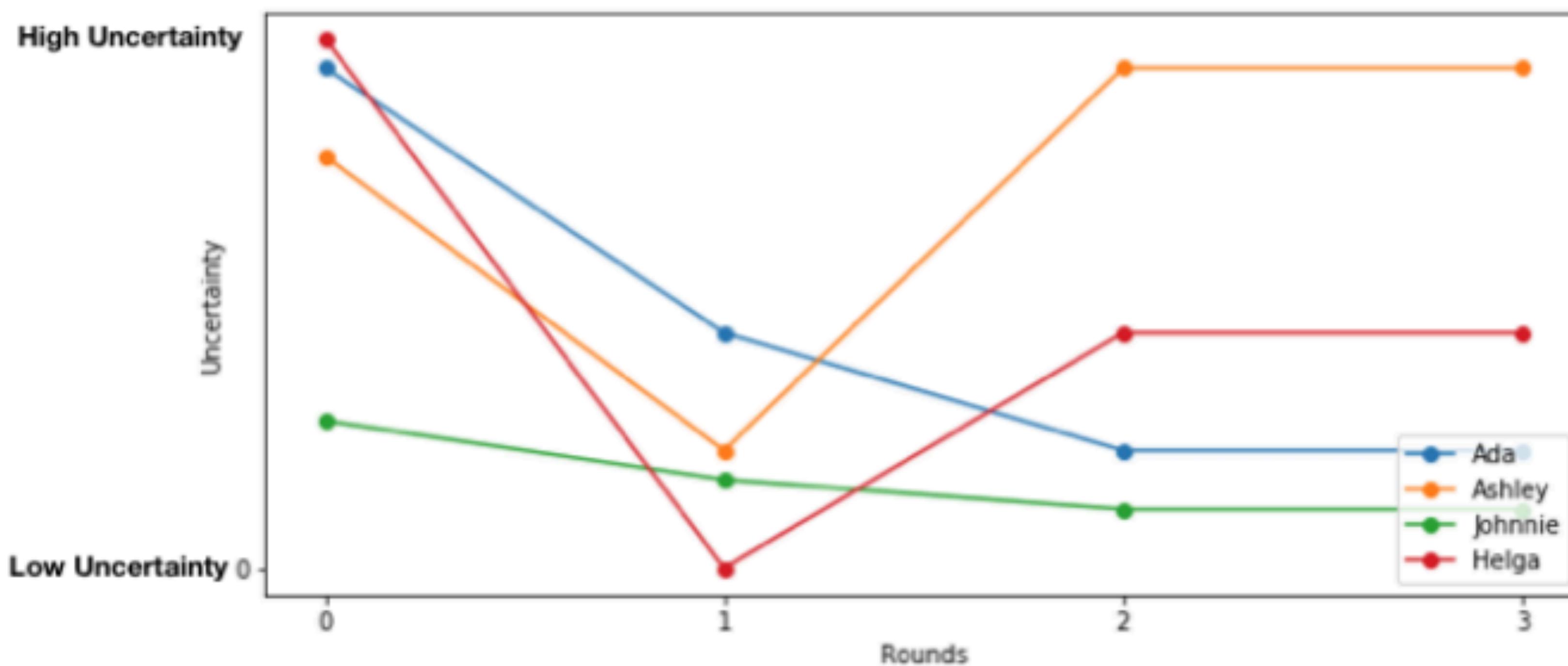
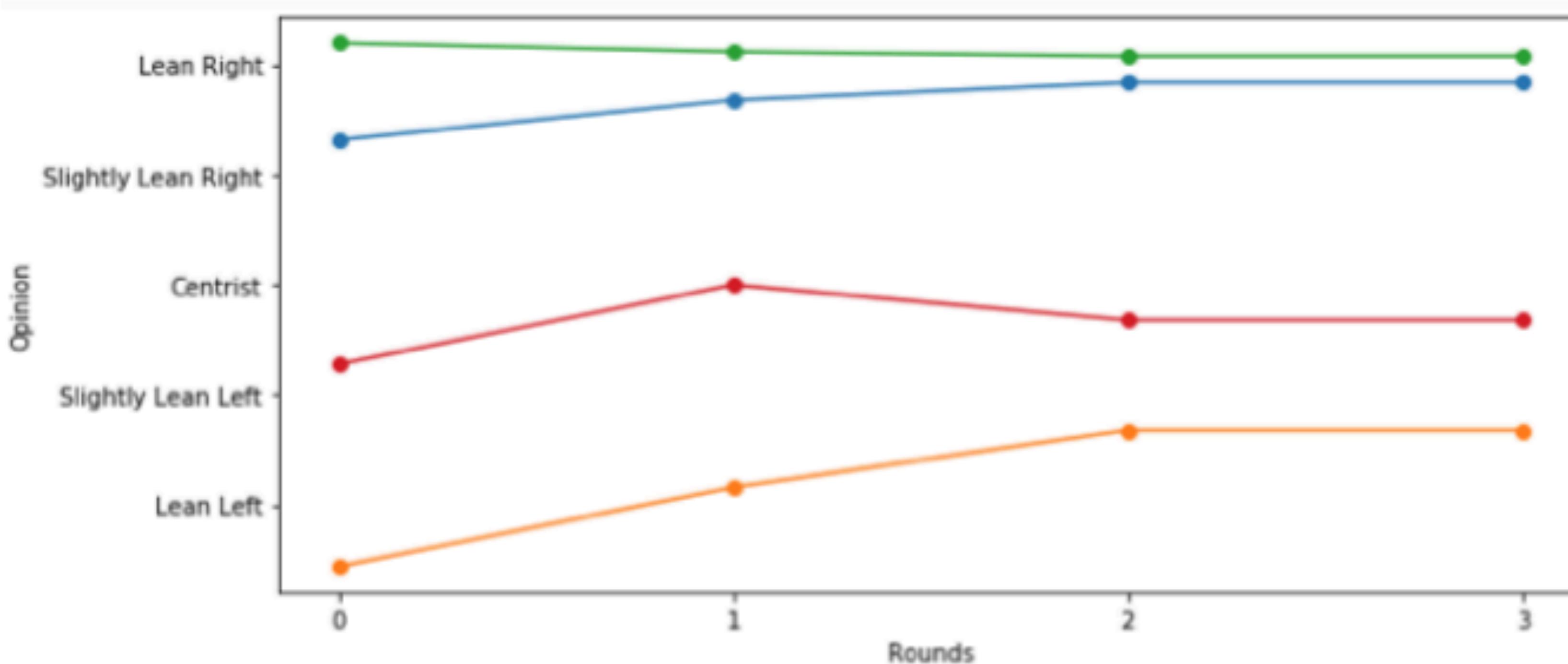
#StandingGround  
#UsedPoliticalAffiliationStereotype  
#IdentifyingSimilarGroups #Believable  
#Expected

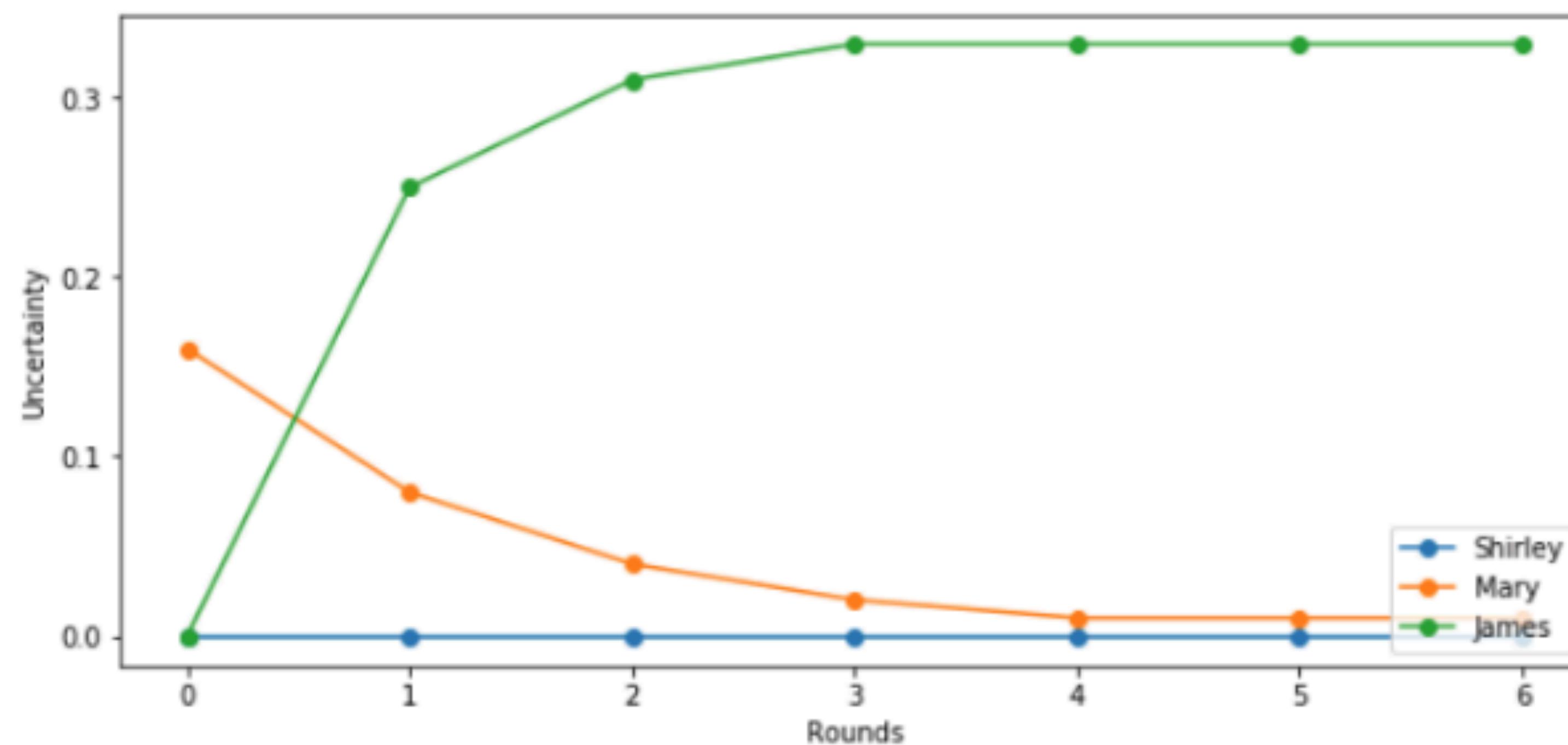
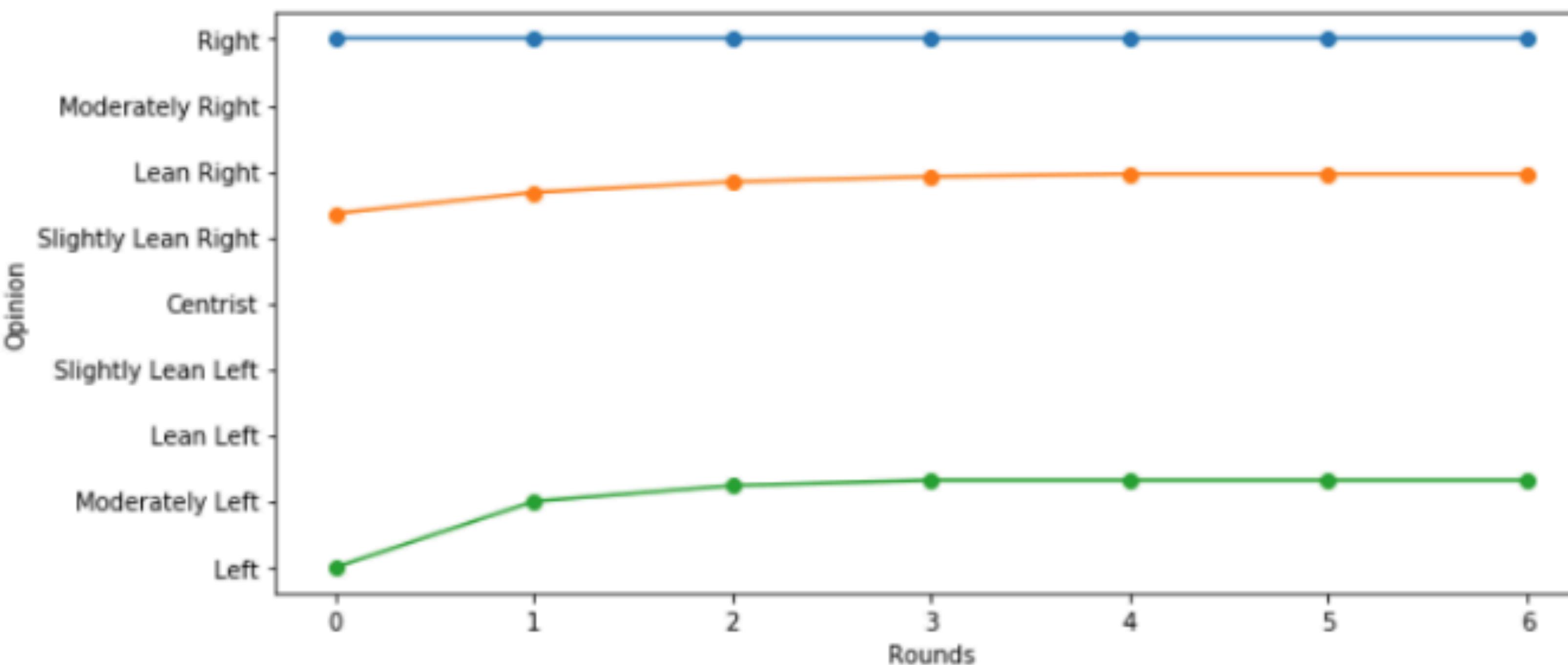
1-14: The way that people's opinions tended to move towards the opinions of those who had similar opinions; causing clusters to slowly emerge.

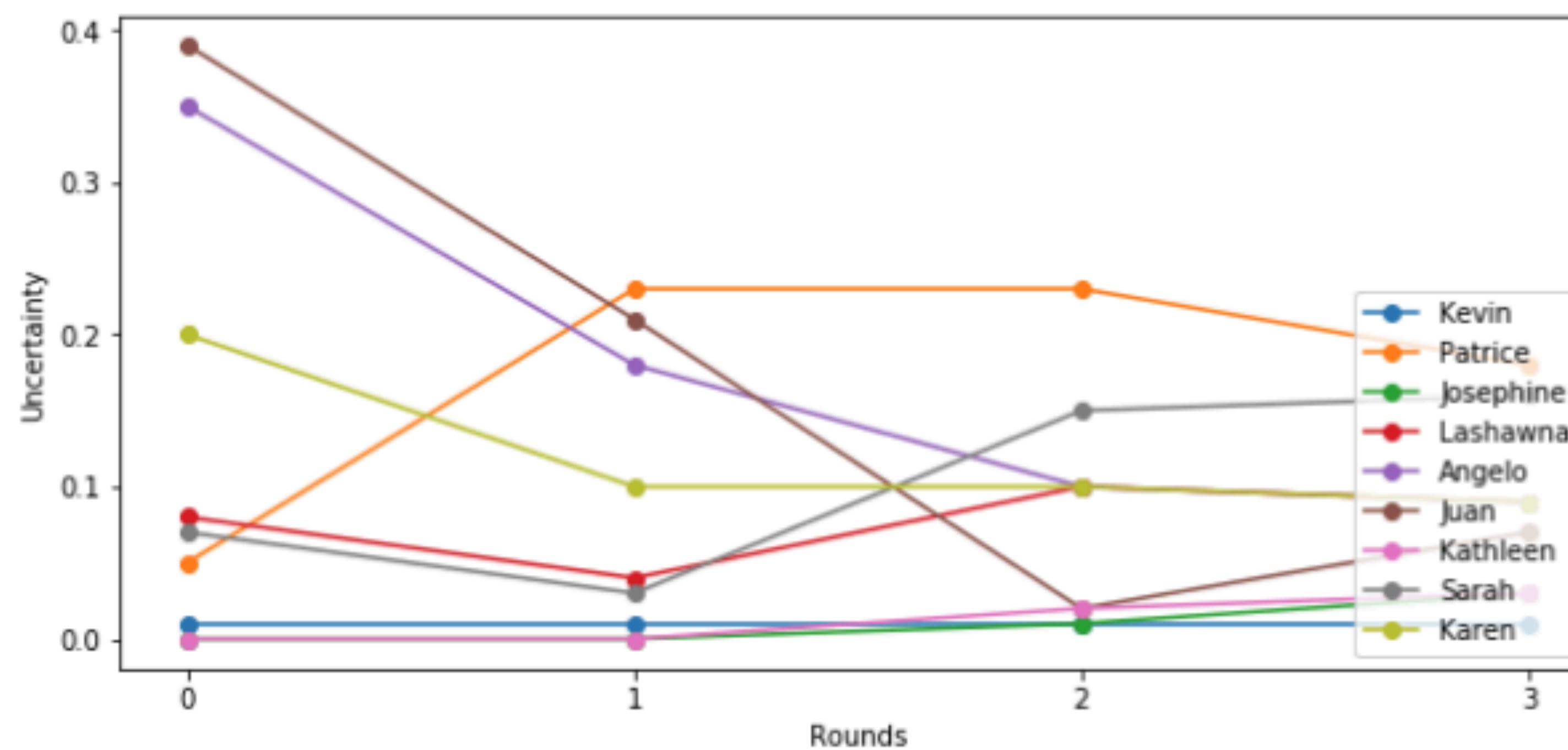
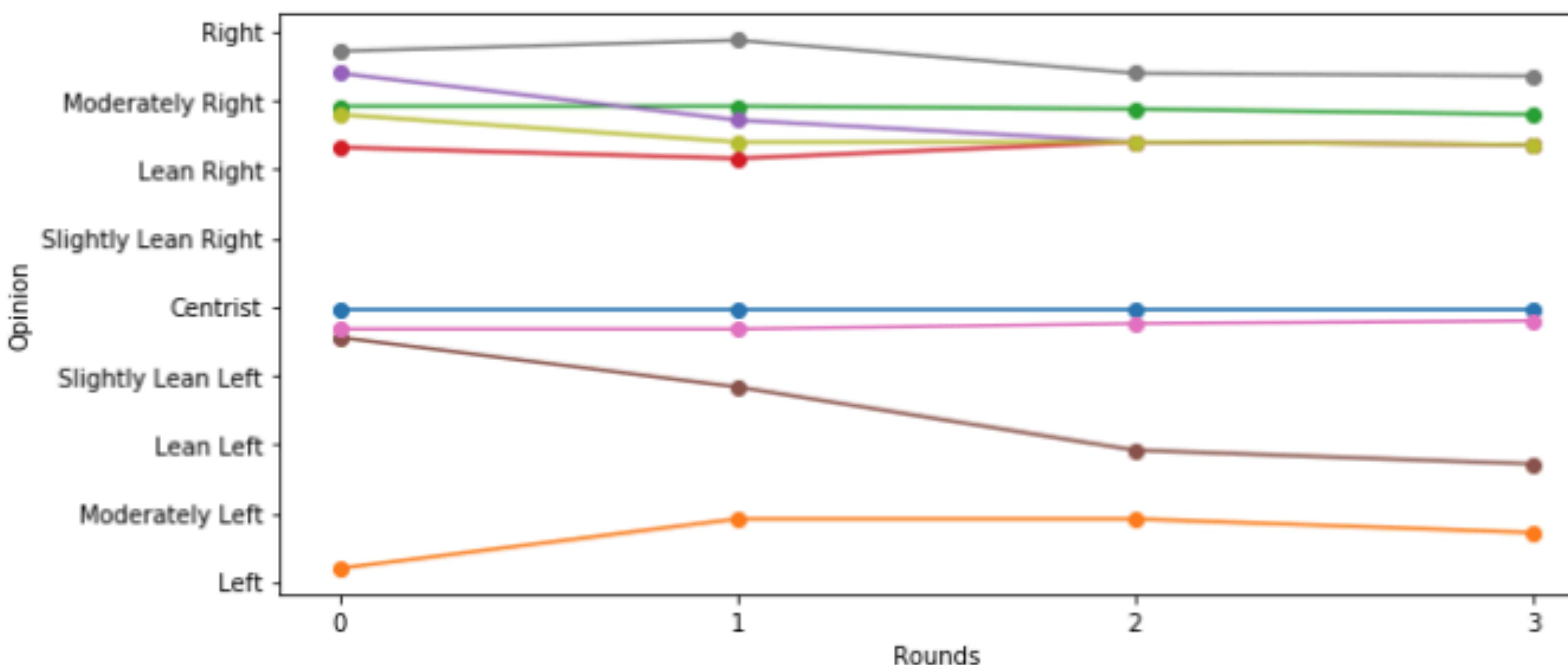
#GroupInfluence #Polarization  
#SimilarViewsConverge #Believable  
#Expected #ClusteringBelievable

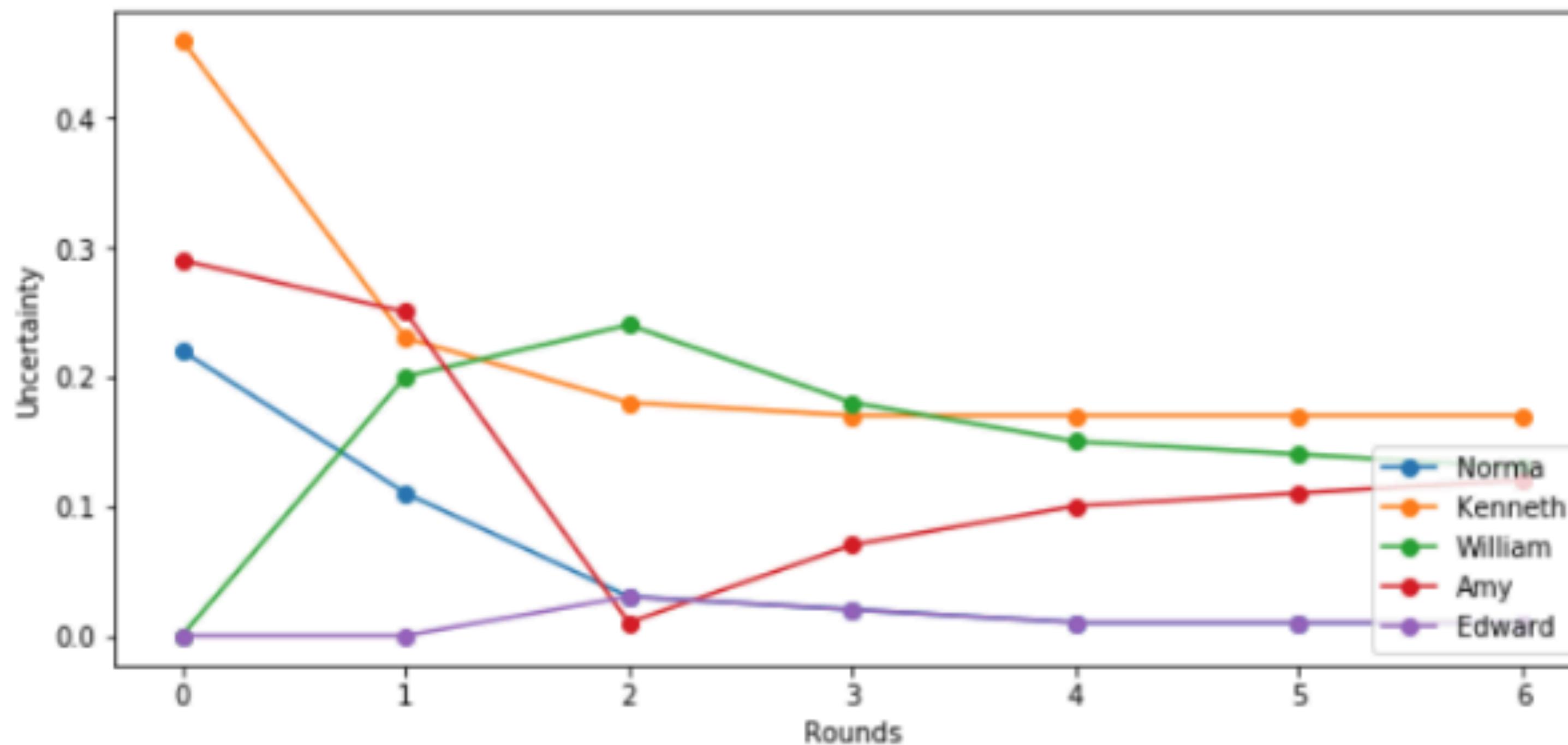
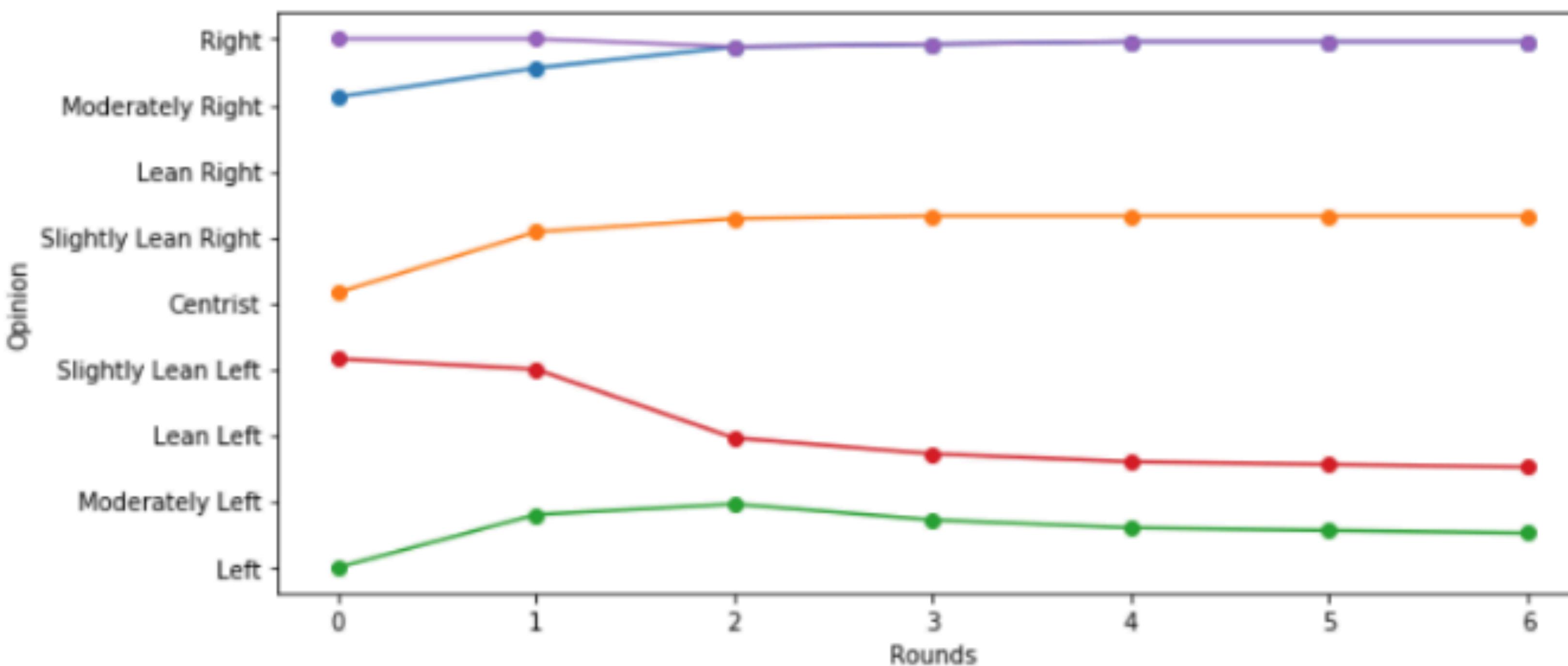
1-16: that there are two groups formed by the two left-of-center people and the two right-of-center people

#Polarization #ClusteringBelievable  
#Expected #Believable









# Choice of Case Study

- Datasets considered: Pro/Con, IMDB, Conference Papers
- The age of political discourse!
  - Founding Father, Benjamin Rush, was convinced – most days, anyway – that there had to be a way to angrily debate the most contentious ideas without ripping the nation apart.

# Features of the AllSides Dataset



**Russia Calls New US Sanctions Illegal**

The Kremlin has condemned as “unacceptable” new sanctions imposed by the Trump administration on Russia over Moscow’s alleged involvement in the poisoning of a former Russian double agent.

**From the Left**

Kremlin calls new Trump sanctions ‘unacceptable’  
ABC News 

**From the Center**

US Sanctions Put Russia on Edge  
Wall Street Journal- Ne 

**From the Right**

Putin had it coming: Trump levies new sanctions against...  
Washington Examiner 

- API accessing corpus of daily news articles
- Grouped by political issues/tags
- Tagged by media bias (source bias, individual bias)

# Features of the AllSides Dataset

News Source	AllSides Media Bias Ranking
New York Daily News	Left
New York Times	Lean Left
Associated Press	Center
Boston Herald	Lean Right
Fox News Editorial	Right

Examples of media bias ratings for various news sources

# Example Discussion

- **Object of Discussion:** Discussion on news article “Room for Debate: Should ‘Birthright Citizenship’ Be Abolished”
- **Source:** NY Times (Bias: Leaning Left)
- **Where:** At work with colleagues
- **Topic:** Immigration
- **Duration:** 11 minutes
- **Number of participants:** 4

# Example Discussion

Discussion for 11 minutes on "Room for Debate: Should 'Birthright Citizenship' Be Abolished?"

Source: NYTimes

Participants: Richard Cain, Ruth Franklin, Vickie Nguyen-Self, Suzanne Sorenson

## Views:

Richard: less uncertain.

Old --> att: -0.5 | op: -0.559 | unc: 0.632

New --> att: -0.5 | op: -0.559 | unc: 0.554

Vickie: change in views, less uncertain.

Old --> att: 1 | op: 0.535 | unc: 0.464

New --> att: 0.948 | op: 0.521 | unc: 0.497

Ruth: no change in views.

att: 0.0 | op: -0.063 | unc: 0.225

att: 0.0 | op: -0.063 | unc: 0.225

Suzanne: no change in views

att: -0.5 | op: -0.559 | unc: 0.059

att: -0.5 | op: -0.559 | unc: 0.059

Left: -1.0 | Leaning Left: -0.5 | Center: 0 | Leaning Right: 0.5 | Right: 1.0