Social Bot Persuasion Techniques in COVID-19 Misinformation

Lynnette Hui Xian Ng
Carnegie Mellon University

Wenqi Zhou
Duquense University

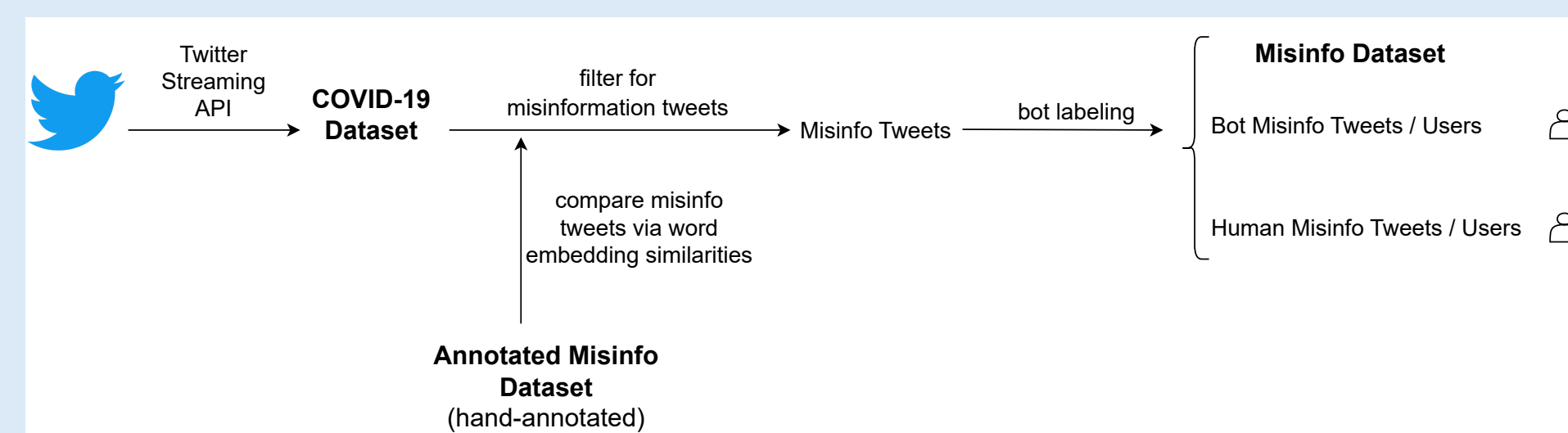Kathleen M. Carley
Carnegie Mellon University

Human Biases are shortcut the brain uses to make sense of uncertainty

RQ1: Can we methodically identify triggers of human biases?
RQ2: What extend did triggers facilitate engagement?

Data

A year's worth of COVID-19 Misinformation data from Twitter
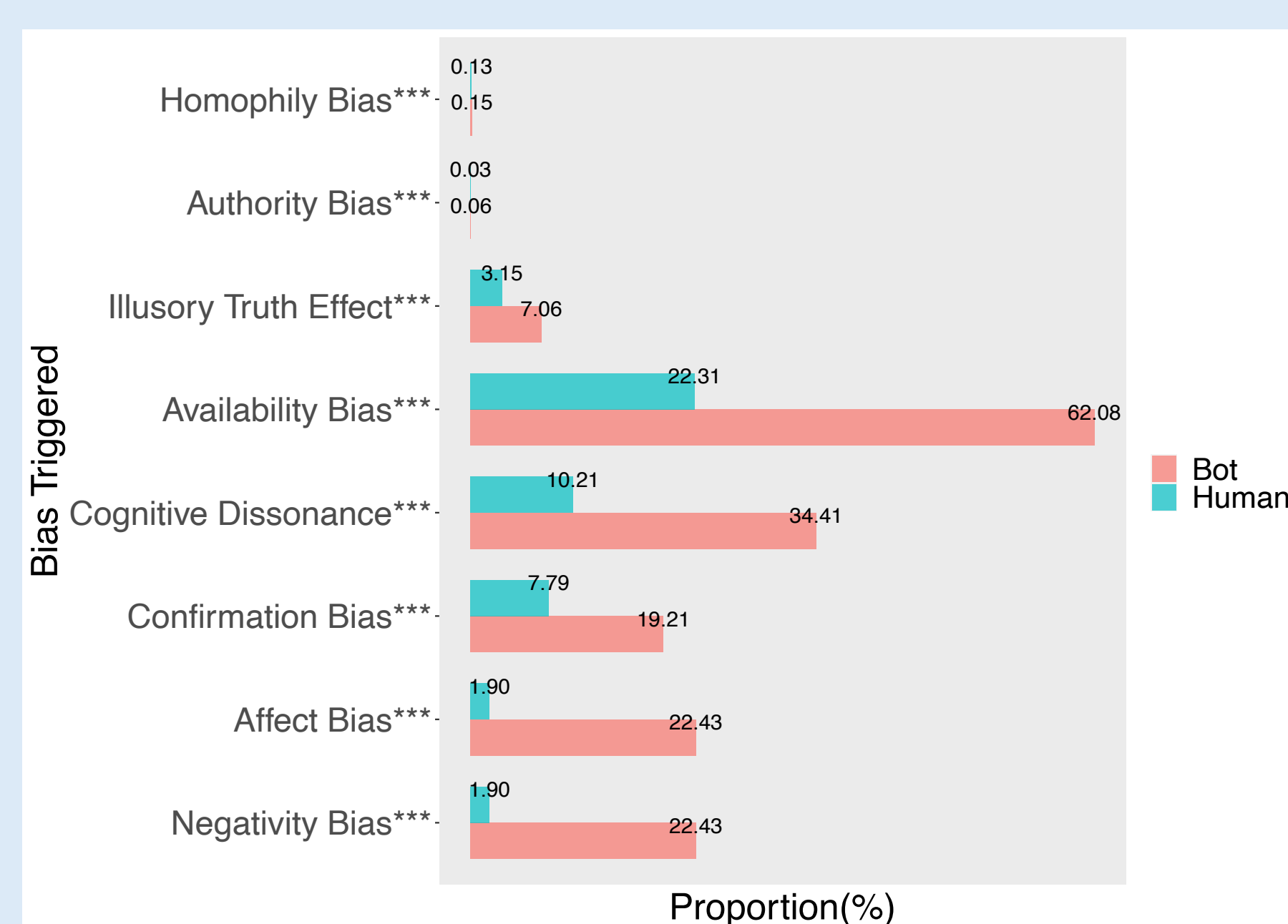

Proportion of biases

Bots trigger more biases than Humans

Majority of Human tweets did not employ biases, but majority of Bot tweets did

Bots are more masterful at forging fake collective belief, changing their opinion positions to blend in

Emotional appeal is a highly used persuasive strategy

Persuasion Techniques

We measure persuasion by presence of bias triggers within the tweets. We derived heuristics to computationally identify triggers at scale after a round of annotations.

Bias triggers, definitions:

| Judgment by Representativeness	The similarity between two groups represent the degree which one is representative of the other
Homophily Bias	Increased acceptance of a message if it is shared to the same community
Authority Bias	Perception that message is credible if it is from reputable sources
Judgment by Availability	An event is judged likely if it is easy to imagine or memorable
Availability Cascade	Collective belief gains more plausibility through increased repetition
Illusory Truth Effect	Tendency to believe in a story more easily when it is repeated multiple times
Affect Bias	Emotions rather than information have a disproportionate effect on judgment
Negativity Bias	Higher belief in negative than positive news
Judgment by Anchoring	Where an anchor is used as an approximation towards the judgment
Cognitive Dissonance	Tendency to avoid having conflicting beliefs and attitudes
Confirmation Bias	Favor information that conforms and strengthen prior beliefs

Co-Occurrence of Biases

(Confirmation Bias, Cognitive Dissonance)
Change stance & remain in stance for a few more tweets
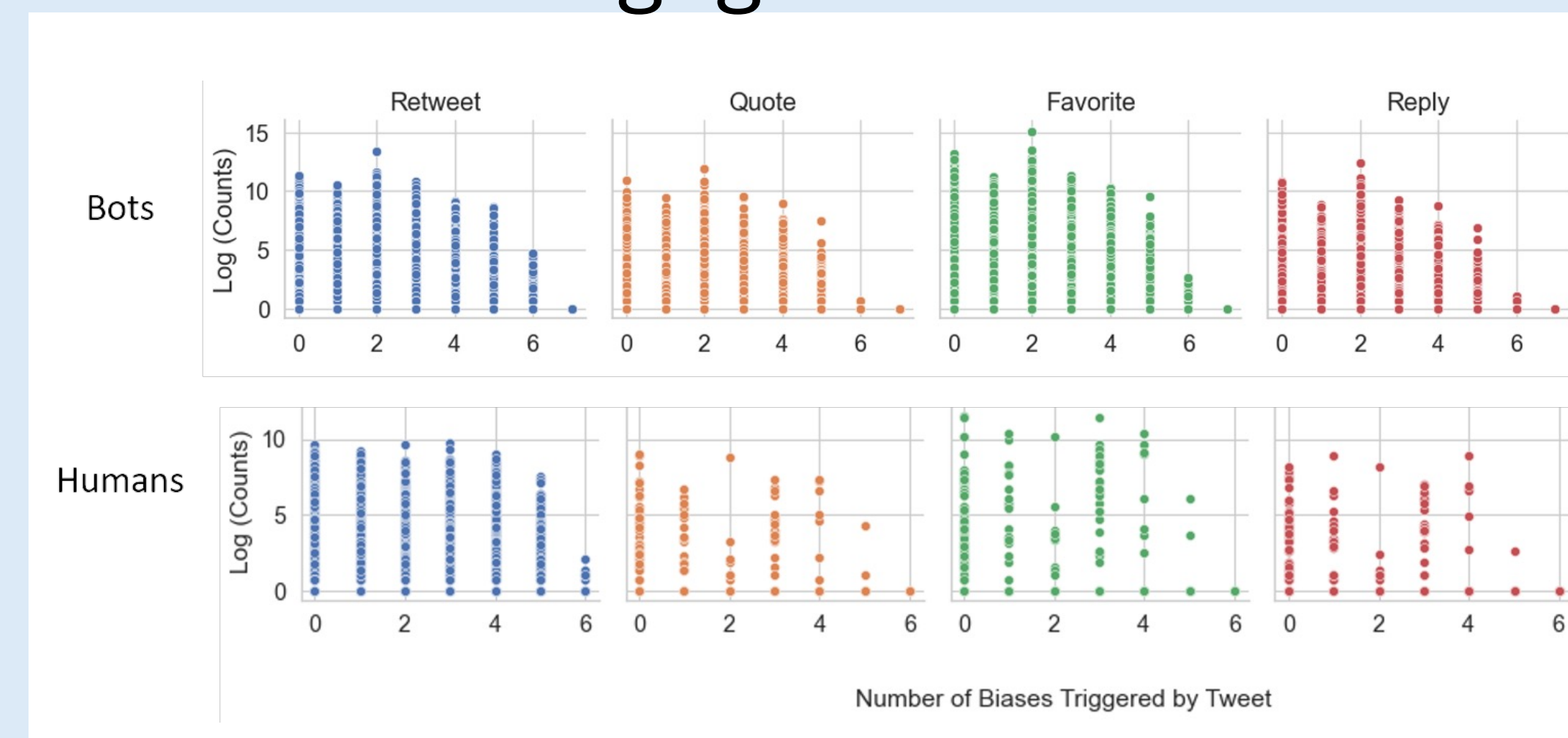
(Affect Bias, Availability Bias)
Emotional tweets more likely to be retweeted
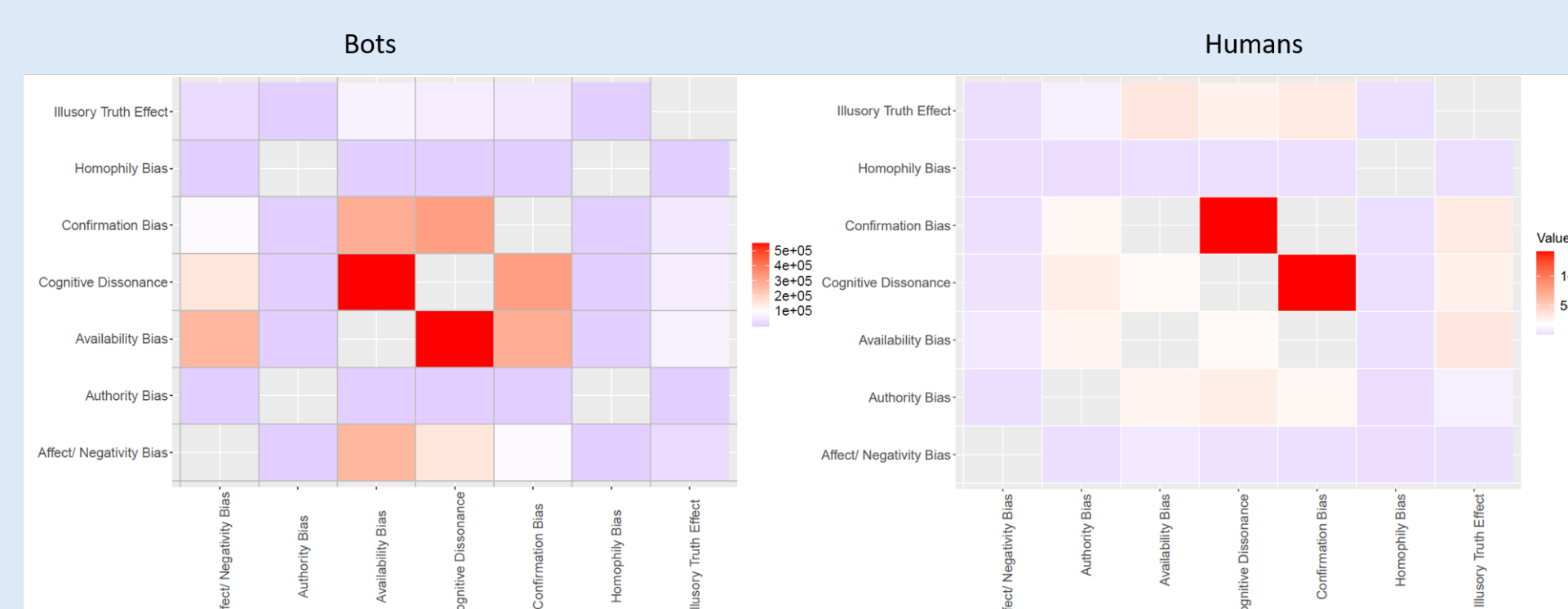
Engagement vs. Bias

Bots have consistent relationship between engagement and number of biases

Decreased engagement after 2 biases

Impact of Bias Triggers on Tweet Engagement

Homophily Bias: Group similarity doesn't quite affect engagement
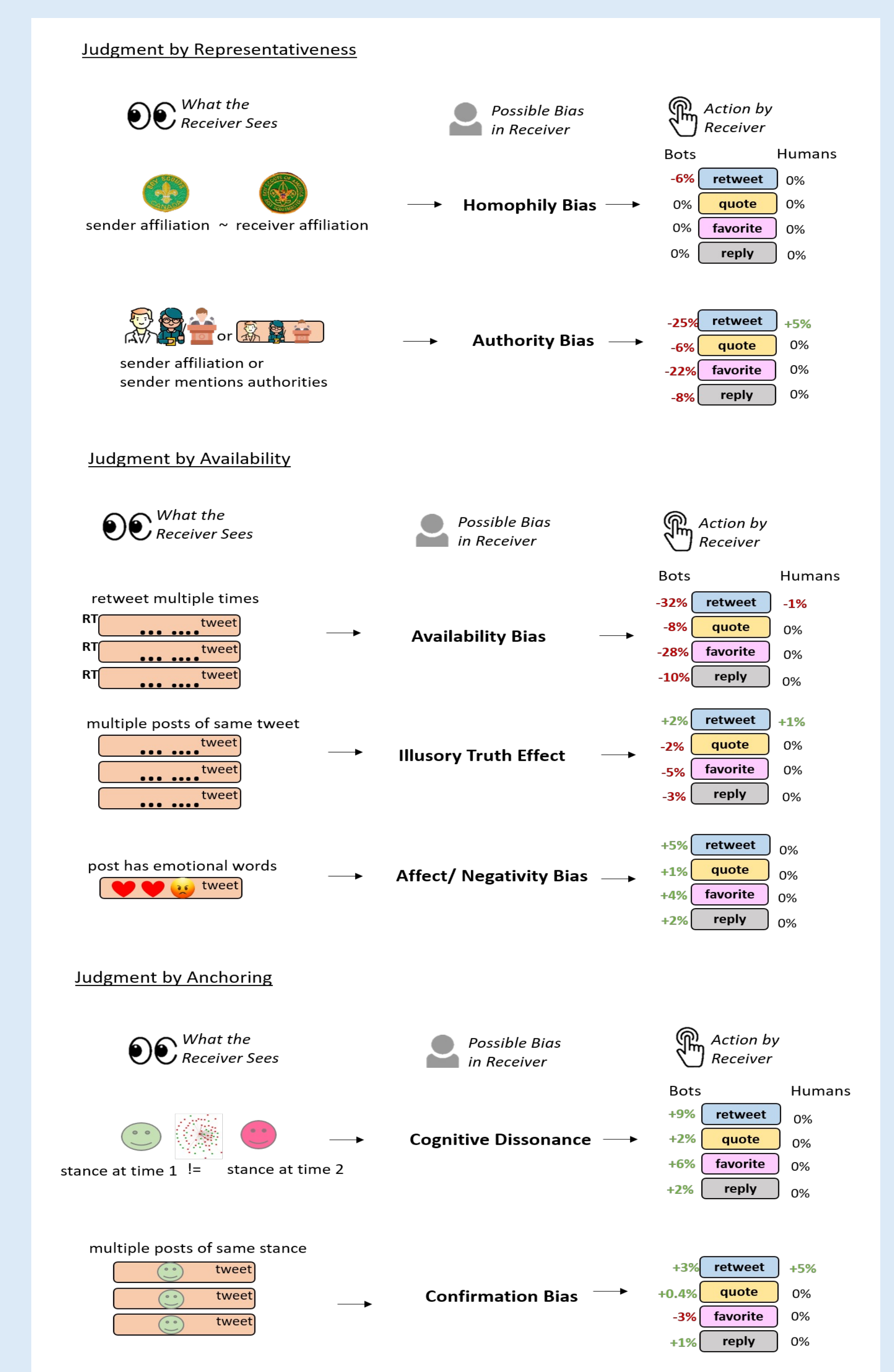
Authority Bias: People often reject explicit authority

Availability Bias & Illusory Truth Effect: Too much information causes cognitive overload; increased repetition repels people

Affect/ Negativity Bias: Emotional appeal increases engagement

Cognitive Dissonance: People are more receptive to those who fit in

Confirmation Bias: People like reassurance

Link to preprint