# BIAS
# Against Asians, Blacks, Hispanics/Latinx, Muslims, and Jews on Twitter

Preliminary results from a collaborative research project with 140 students at Indiana University
December 2021

# Executive Summary

We looked at conversations on Twitter about Asians, Blacks, Latinos, Muslims, and Jews by annotating and classifying 10,000 tweets.

We created representative samples of tweets with these keywords for two periods.

- January to June 2021 for the keywords "Asians, Blacks, and Latinos, and Muslims" and January to April for the keyword "Jews."
- July to September 2021 for the keywords "Asians, Blacks, and Latinos, and Muslims" and May to August for the keyword "Jews."

➢ **Annotators saw most biased tweets in conversations about Jews and black people, followed by conversations about Muslims, Asians, and Hispanics/Latinos.**

➢ **Conversations about Jews from May to August 2021 had the highest number of biased tweets across all samples and showed a strong increase from the previous period. This indicates that the Israeli-Palestinian conflict triggers not only criticism of Israel but also bias against Jews. Positive sentiments declined, and negative sentiments increased in the period of May to August.**

➢ **However, conversations about Jews also had the highest number of tweets that were calling out bias.**

# Methods

**Raw data:** query results from a 10 % database of all tweets (IU's Observatory on Social Media) with the keywords "Asians, Blacks, Latinos, Muslims, and Jews" from January to September 2021.

**Representative samples** of about 500 tweets for each of the five keywords and for two time periods, January-June 2021 (January to April for the keywords "Jews") and July-September 2021 (May to August for the keywords "Jews").

**Annotation** of the representative samples (in subsamples of 100) by contributing students. We classified each tweet in the samples, looking at bias, sentiments, and calling out bias by using a specially designed annotation portal, see annotationportal.com/.

Annotators used a scale from 1 to 5 (confident not biased, probably not biased, don't know, probably biased, confident biased). Bias definitions for each category can be found in the annex. All annotators (140 participating students) annotated 200 tweets or more.
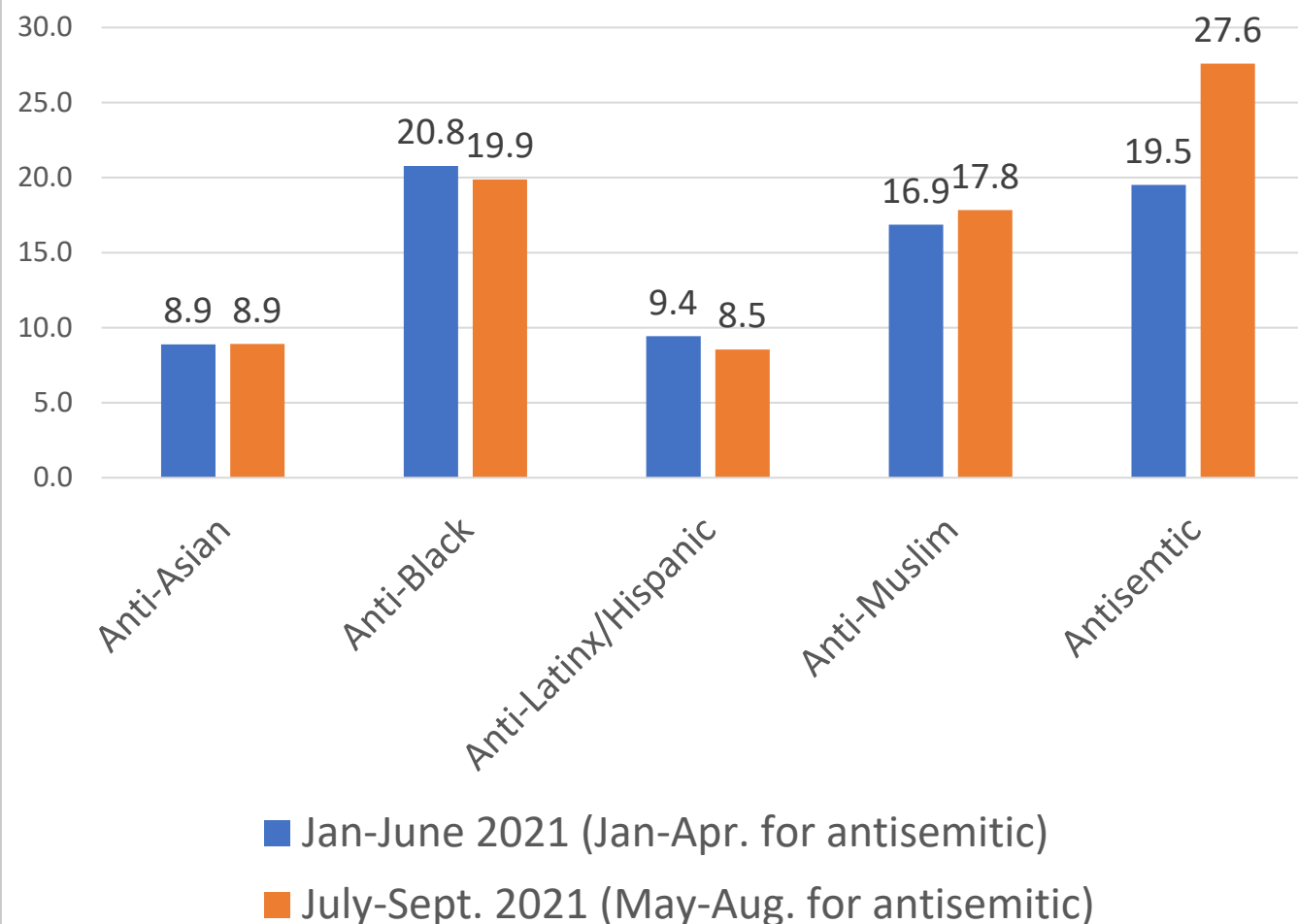
# Preliminary Results

Biased tweets in conversations on Twitter about Asians, blacks, Latinos, Muslims, and Jews (representative samples of tweets with those keywords).

Caution: The results show the average percentage of tweets the annotators classified as biased (confident and probably). Disagreements on the classification of individual tweets are not considered here. For an analysis by Median and Means of each tweet that takes disagreements into account, see further below.
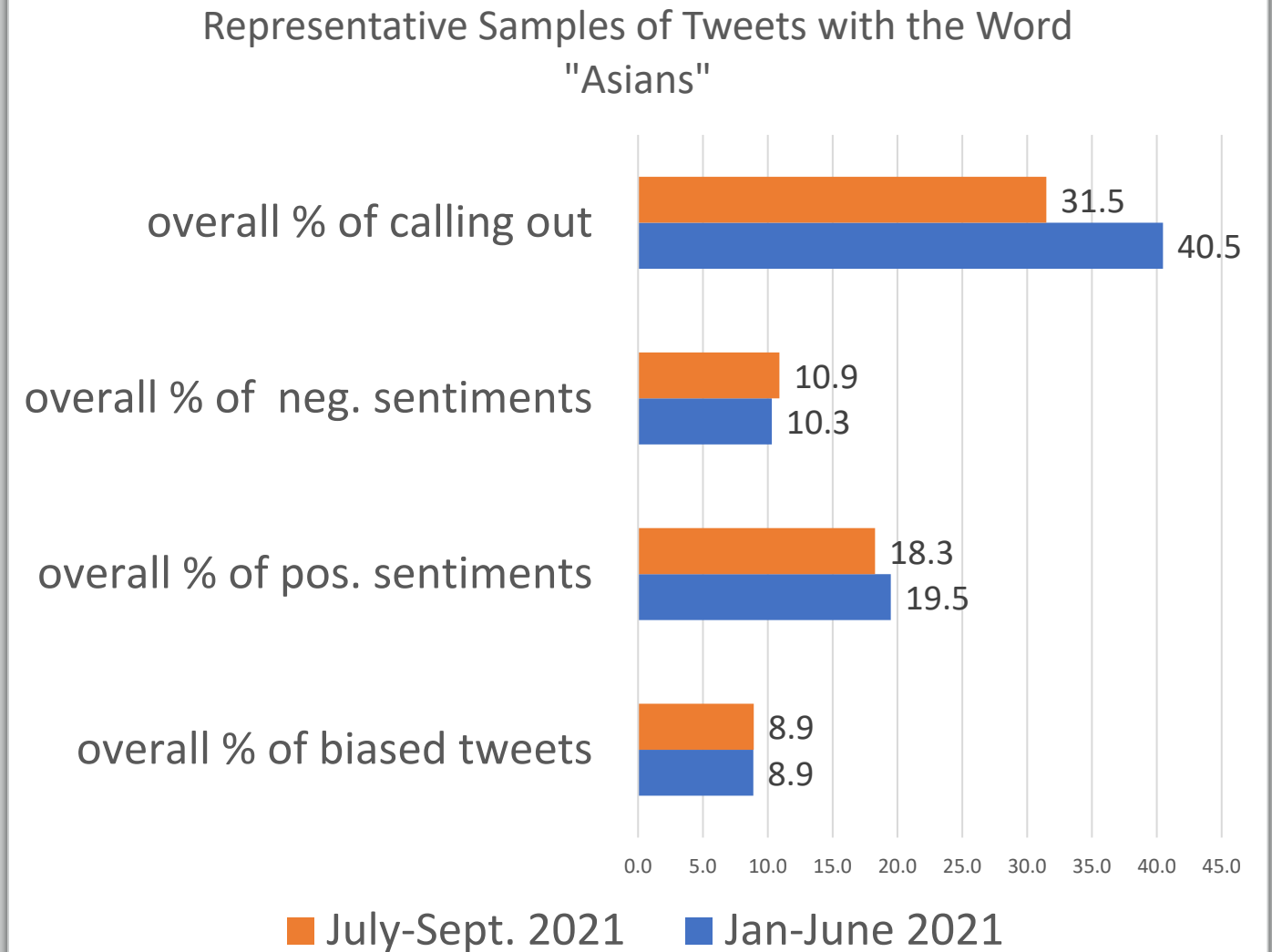
Percentage of Biased Tweets on Average Across Annotators in Conversations About Asians, Blacks, Latinos, Muslims, and Jews

| Category | Jan-June 2021 (Jan-Apr. for antisemitic) | July-Sept. 2021 (May-Aug. for antisemitic) |
|---|---|---|
| Anti-Asian | 8.9 | 8.9 |
| Anti-Black | 20.8 | 19.9 |
| Anti-Latinx/Hispanic | 9.4 | 8.5 |
| Anti-Muslim | 16.9 | 17.8 |
| Antisemtic | 19.5 | 27.6 |

■ Jan-June 2021 (Jan-Apr. for antisemitic)
■ July-Sept. 2021 (May-Aug. for antisemitic)

# Annotation Results of the Samples with the Keyword "Asians"
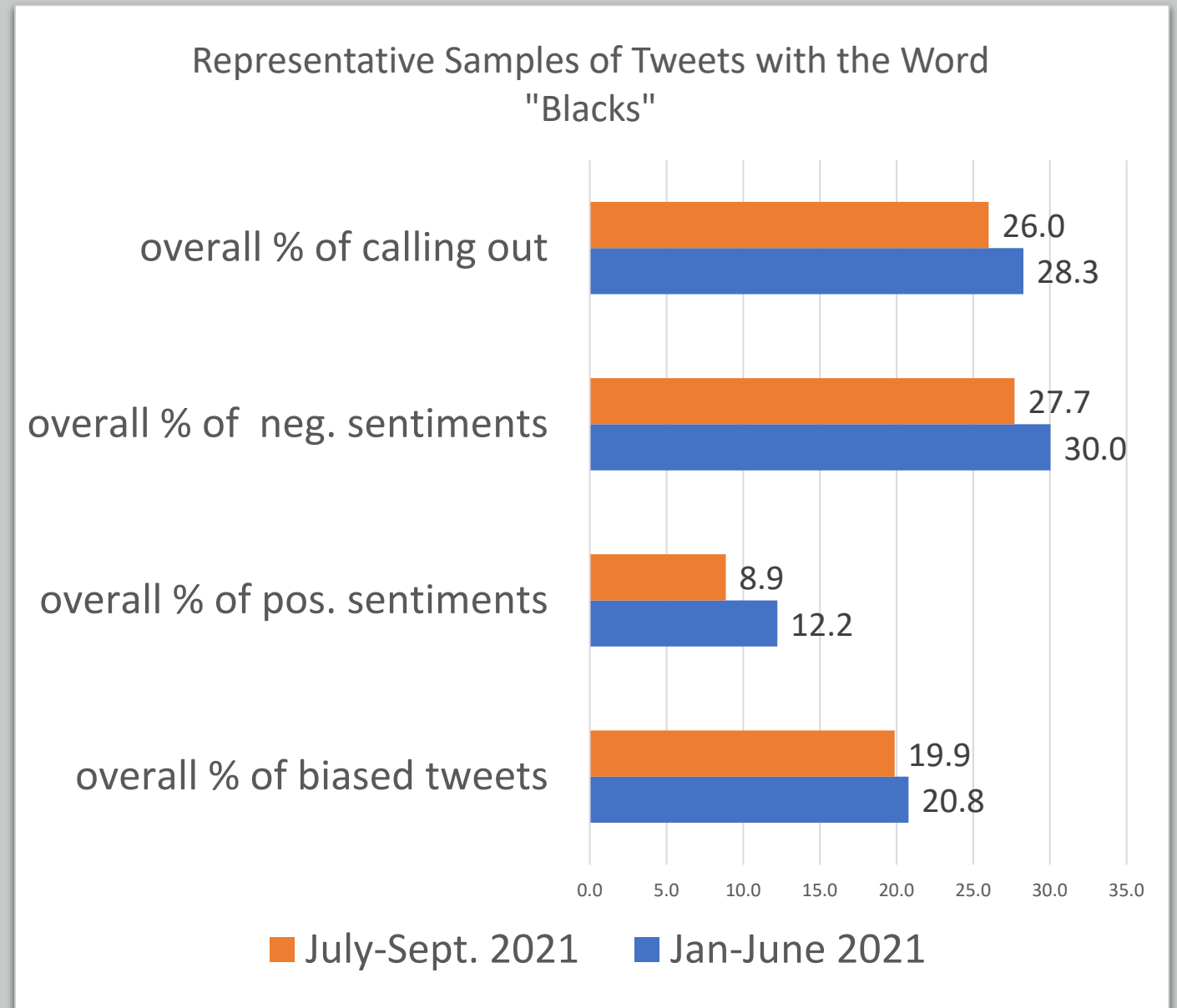
On average, annotators classified
- about a third of the tweets as calling out bias;
- more tweets as showing positive than negative sentiments;
- 8.9 % of the tweets as biased in both time periods.

Representative Samples of Tweets with the Word "Asians"



overall % of calling out — July-Sept. 2021: 31.5, Jan-June 2021: 40.5
overall % of neg. sentiments — July-Sept. 2021: 10.9, Jan-June 2021: 10.3
overall % of pos. sentiments — July-Sept. 2021: 18.3, Jan-June 2021: 19.5
overall % of biased tweets — July-Sept. 2021: 8.9, Jan-June 2021: 8.9

■ July-Sept. 2021   ■ Jan-June 2021

# Annotation Results of the Samples with the Keyword "blacks"

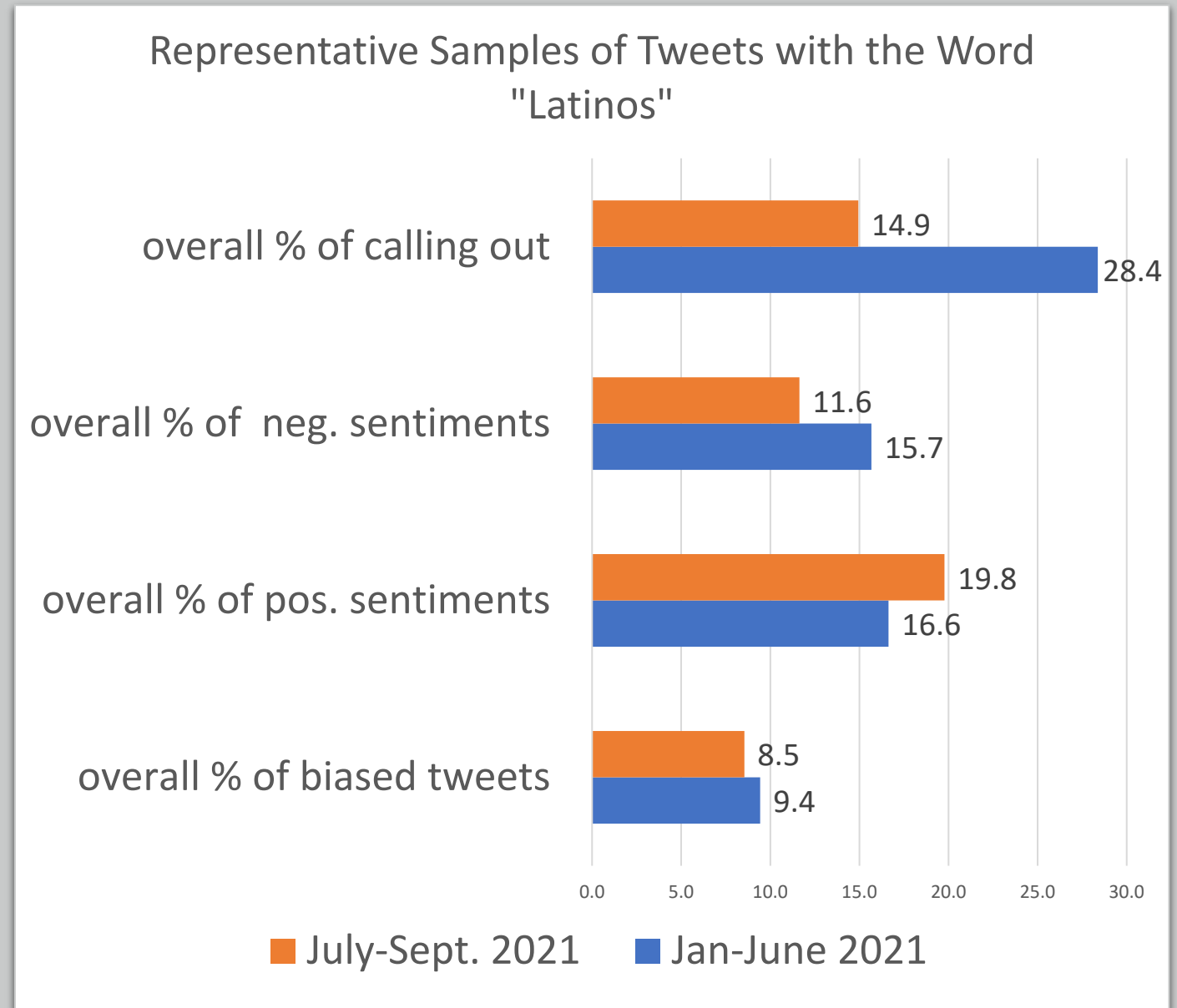On average, annotators classified

- 26/ 28 % of the tweets as calling out bias;
- more tweets as showing negative than positive sentiments;
- 1 in 5 tweets as biased in both time periods.

Representative Samples of Tweets with the Word "Blacks"

| | July-Sept. 2021 | Jan-June 2021 |
|---|---|---|
| overall % of calling out | 26.0 | 28.3 |
| overall % of neg. sentiments | 27.7 | 30.0 |
| overall % of pos. sentiments | 8.9 | 12.2 |
| overall % of biased tweets | 19.9 | 20.8 |

# Annotation Results of the Samples with the Keyword "Latinos"

On average, annotators classified

• 28 % of the tweets as calling out bias in the first time period from January to June but only 15 % from July to September;

• more tweets as showing positive than negative sentiments;

• about 9 % of the tweets as biased in both time periods.

Representative Samples of Tweets with the Word "Latinos"

overall % of calling out
14.9
28.4

overall % of neg. sentiments
11.6
15.7

overall % of pos. sentiments
19.8
16.6

overall % of biased tweets
8.5
9.4

0.0   5.0   10.0   15.0   20.0   25.0   30.0
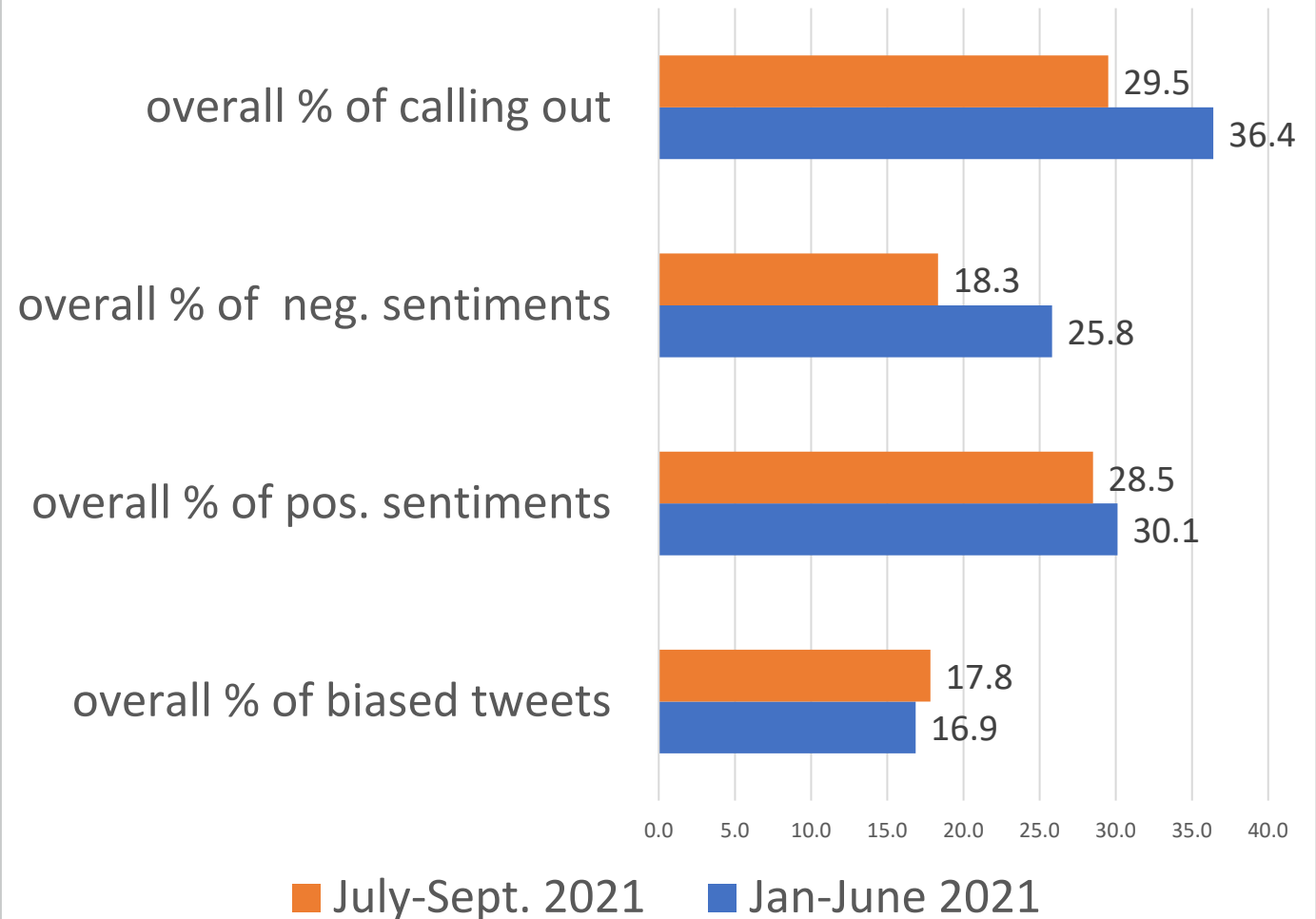
■ July-Sept. 2021   ■ Jan-June 2021

# Annotation Results of the Samples with the Keyword "Muslims"

On average, annotators classified

- About a third of the tweets as calling out bias;
- more tweets as showing positive than negative sentiments;
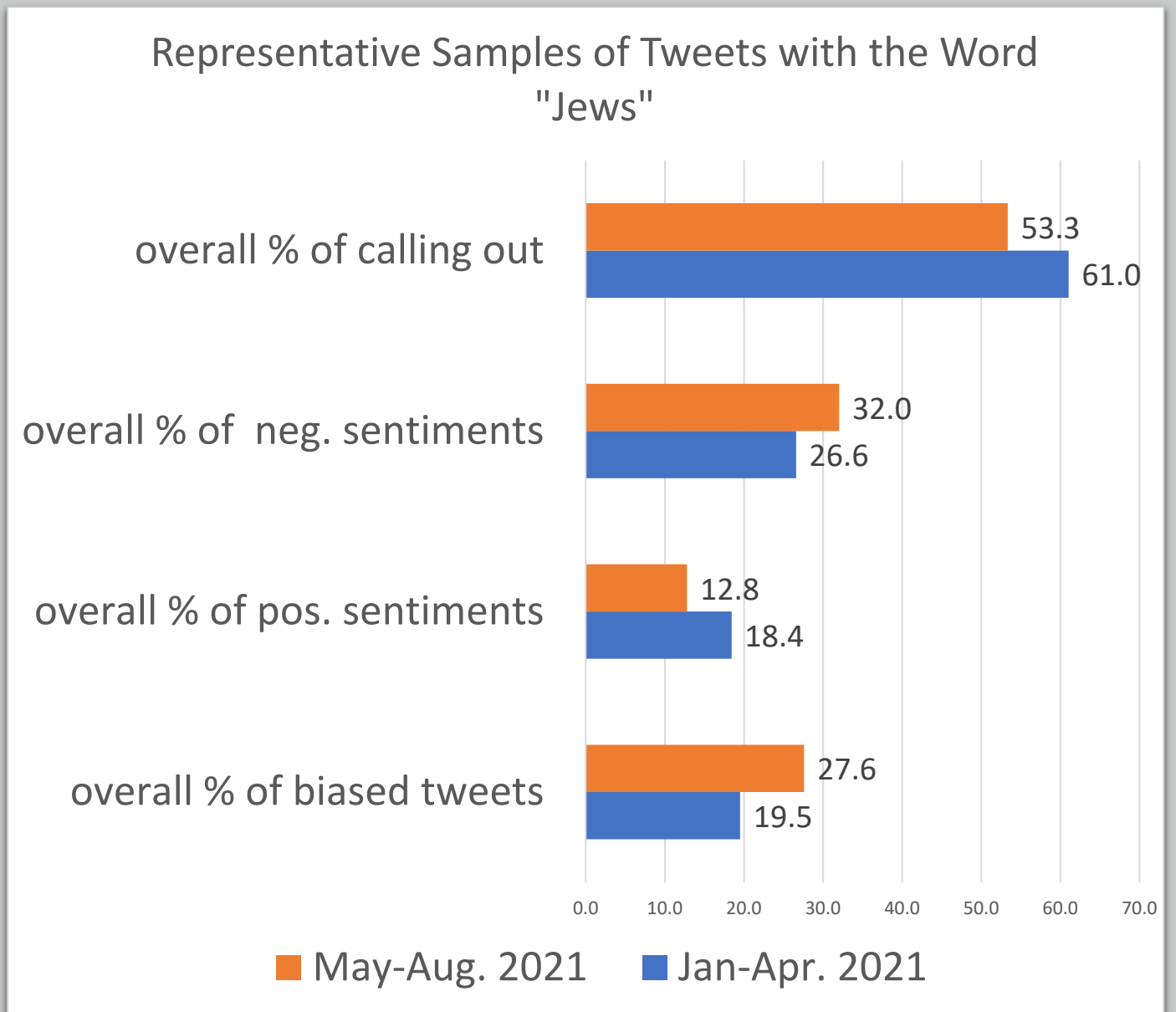- about 17 % of the tweets as biased in both time periods.

Representative Samples of Tweets with the Word "Muslims"

| | July-Sept. 2021 | Jan-June 2021 |
|---|---|---|
| overall % of calling out | 29.5 | 36.4 |
| overall % of neg. sentiments | 18.3 | 25.8 |
| overall % of pos. sentiments | 28.5 | 30.1 |
| overall % of biased tweets | 17.8 | 16.9 |

# Annotation Results of the Samples with the Keyword "Jews"

On average, annotators classified

• more than half of the tweets as calling out antisemitism;

• more tweets as showing negative than positive sentiments;

• More tweets as antisemitic in the period May to August 2021 than in the period January to April, possibly related to the war between Hamas and Israel in May 2021.



Representative Samples of Tweets with the Word "Jews"

overall % of calling out — 53.3 / 61.0

overall % of neg. sentiments — 32.0 / 26.6

overall % of pos. sentiments — 12.8 / 18.4

overall % of biased tweets — 27.6 / 19.5

■ May-Aug. 2021   ■ Jan-Apr. 2021

# Bias Percentage By Mean and Median for Each Tweet

The mean (average) of bias classification for each tweet is found by adding the numbers from each annotator (1=confident not biased, 2=probably not biased, 3=don't know, 4=probably biased, 5=confident biased) and then dividing by the number of annotators who classified the tweet.

The median is the middle value when the classification of a tweet is ordered from least to greatest.

The following shows the results for two subsets
- Part 5 of representative sample of the time period January to June 2021 with the keyword "blacks" (99 tweets) and
- Part 5 of representative sample of the time period July to September 2021 with the keyword "blacks" (100 tweets).
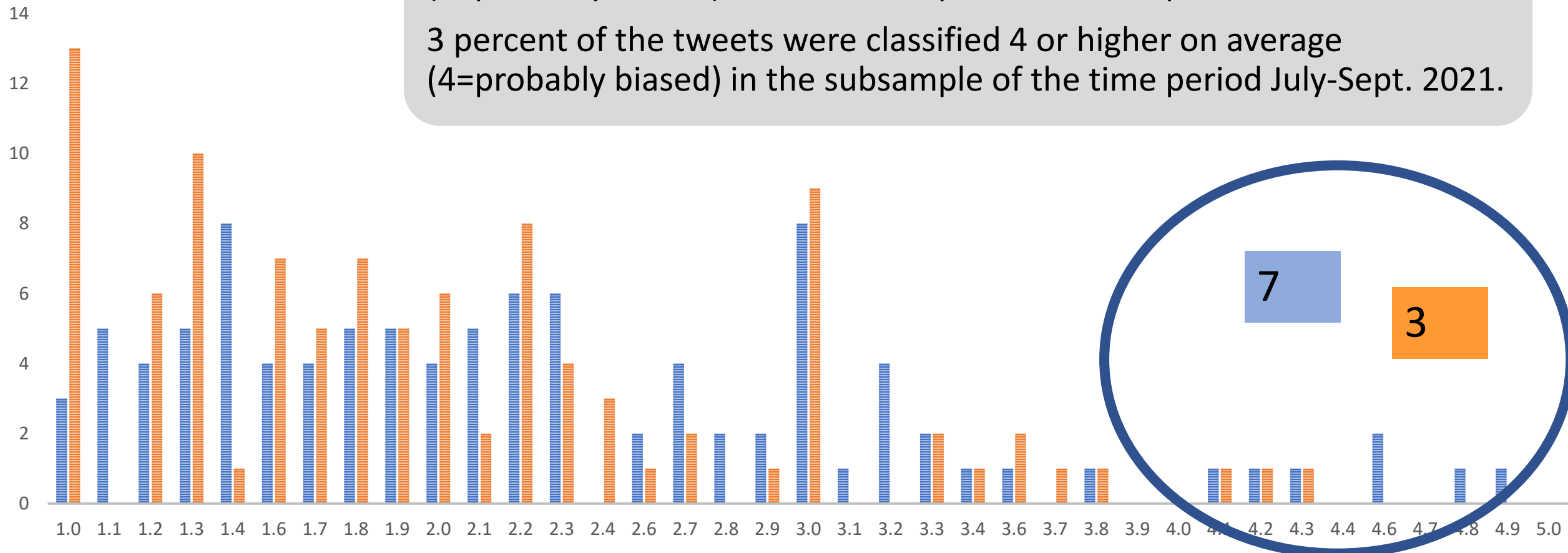
ANTI-BLACK BIAS? PART5 DISTRIBUTION BY MEAN (9 ANNOTATORS)

Jan-June    Jul-Sept.

7 percent of the tweets were classified 4 or higher on average (4=probably biased) in the subsample of the time period Jan-June 2021.

3 percent of the tweets were classified 4 or higher on average (4=probably biased) in the subsample of the time period July-Sept. 2021.
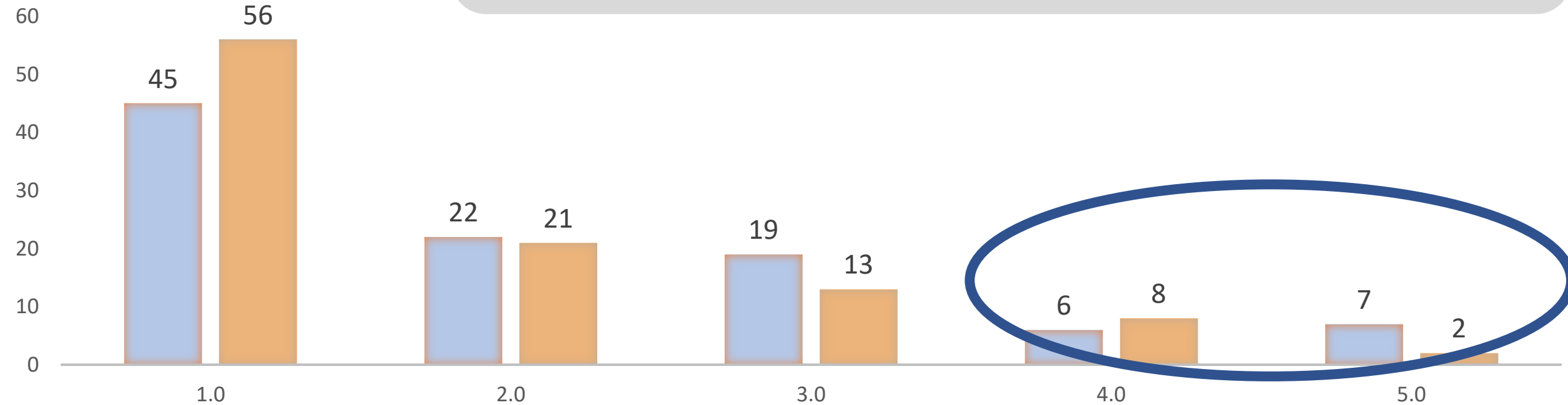
7    3

1.0 = all annotators classified the tweet as conf. not biased

5= all conf. biased

# ANTI-BLACK BIASED? PART5 DISTRIBUTION BY MEDIAN (9 ANNOTATORS)

■ Jan-June   ■ Jul-Sept.

The middle value of 13 percent of the tweets was 4 or higher (4=probably biased) in the subsample of the time period Jan-June 2021. The middle value of 10 percent of the tweets was 4 or higher (4=probably biased) in the subsample of the time period July-Sept. 2021.



1= median (middle value) confident not biased

5= median conf. biased

# The exact percentage of tweets that are classified as biased depends on the method of analysis.

## Percentage of Biased Tweets by Average, Median, and Mean

| Subsample part 5 of representative sample with the keyword "blacks" | Jan-Jun. | Jul-Sept. |
|---|---|---|
| Average percentage of biased tweets | 20.6 % | 15.9 % |
| % of tweets classified as biased by median (middle value) | 13 % | 10 % |
| % of tweets classified as biased by mean (average) across annotators | 7 % | 3 % |

# Contributors

Elisha S. Breton and Walker Rhea (teaching associates and graduate students at the Borns Jewish Studies Program, IU).
Katharina Soemer (graduate student at Bremen University and affiliate at IU).
Students of the class "White Supremacism and Antisemitism on Social Media." (Fall 2021 by Gunther Jikeli): Alexander Ake, Jacob Alarcon, Julia Amodeo, Carlee Anderson, Emma Arganoff, Allie Barrett, Samuel Berebitsky, Josie Bernbom, Peyton Bietry, Hunter Birsfield, Christopher Bradberry, Reese Bricker, Mairead Brown, Bridget Bruni, Alyson Bunn, Jacquelyn Butler, Alex Butrum-Griffith, Christine Campbell, Myah Clayborn, Dresden Cogan, Rachel Cohen, Dayna Cohn, Isabelle Conner, Alex Cooper, Desiree Cousin, Megan Cunningham, Nicole Daniel, Kelsey Dennehy, Linh Dinh Tran, Bryan Doan, Grace Donahue, Lucie Downing, Avery Edwards, Emily Fera, Andrew Filler, Anna Filoso, Grant Fisher, Kyle Flick, Kim Gabriel, Jayla Garnett, Sydney Geis, Keeton Gibson, Rachel Gilchrist, Cassie Golden, Emely Gonzalez, Maya Gross, Kaitlin Harris, Ana Havens, Maya Hogan, Alexandria Hopper, Hannah Hunter, Sophia Jackson, Kaili James, Lauren Jarboe, Abigail Johnson, Kylie Johnson, Sonny Jones, Hannah Katinsky, Kaylyn Kennedy, Abby Kickbush, Maggie Killian, Molly Killinger, Noah Kintner, Kadzumi Komiyama, Billee Krupnick, Madison Landsberg, Caitlyn Leraas, Jaren Linowes, Dawn Lipperd, Yuxiao Liu, Heidi Livingston, Ly Luc, Isabelle Lynch, Alek Maasik, Hailey Mangelsdorf, Nathaniel Mathew, Riley McCoy, Reese Millspaugh, Emma Milton, Kendall Mirise, Lauren Mondello, Rachel Morris, Kennedy Nash, Sophia Newman, Sophie Nix, Hannah Ogden, Grace Pace, Erin Panizzi, Abby Peterson, Caroline Plumleigh, Kayden Poindexter, Aidan Robinson, Alicia Robinson, Juliana Robinson, Karlie Rogers, Abby Rose, Lucy Rosen, Jeremy Rosenblum, Levi Roughley, Nicole Russnak, Marnie Sara, Srishti Sarkar, Malia Scheele, Eleanor Scheper, Abby Schwartz, Olivia Schwartz, Ally Seifried, Annaya Shelton, Hannah Smith, Luke Snyder, Madison Snyder, Salwa Spir, Sammie Squire, Mary Stafford, Luke Steele, Olivia Sterba, Brooklyn Stewart, Jack Stewart, Molly Stiles, Brooke Swanson, Sam Syroney, Aaron Toland, Dylan Wadsworth, Carly Waite, Eliana Wakefield, Hannah Warren, Sierra Weiks, Grace Weiler, Ellie Weinberg, Jourdan Weinzapfel, Kylie Wexler, Mary Etta Williams, Jake Zechman, Tali Ziner.

# ANNEX Definitions

- **Anti-Asian Bias**
- **Anti-Black Bias**
- **Anti-Muslim Bias**
- **Anti-Hispanic/Latino Bias**
- **Antisemitism**

Anti-Asian bias is a certain perception of Asian people, which may be expressed as racial degradation or hatred toward Asians. Rhetorical and physical manifestations of anti-Asian bias are directed toward Asians or perceived Asian individuals and/or their property or toward Asian community institutions. Contemporary examples of anti-Asian bias in public life, the media, schools, the workplace, and in the religious sphere could, taking into account the overall context, include, but are not limited to the examples below.

This tweet is anti-Asian because it
1. uses an anti-Asian slur in a derogatory way or expresses direct hatred or threats against Asians.
2. blames Asians or Chinese Americans for COVID.
3. alleges that Chinese Americans or Asian Americans are a threat to American culture.
4. is a derogatory statement generally associating Asians with repulsive food culture (e.g. eating cats and dogs).
5. makes derogatory claims based on physical attributes (e.g. skinny, "squinted eyes").
6. alleges that Asian women in general are opportunistic sexual beings or predatory "gold diggers."
7. describes Asian women in general as "submissive," "promiscuous," or exotic seductresses.
8. alleges that Asian men are asexual or that they are predators of white women.
9. alleges that Asians are bad drivers.
10. draws on other anti-Asian stereotypes (name in comments).

# ANNEX
## Anti-Black Bias

Anti-Black bias is a certain perception of black people or African Americans, which may be expressed as racial degradation or hatred toward Blacks or African Americans. Rhetorical and physical manifestations of anti-Black bias are directed toward black individuals and/or their property or toward black community institutions. Contemporary examples of anti-Black bias in public life, the media, schools, the workplace, and in the religious sphere could, taking into account the overall context, include, but are not limited to the examples below.

This tweet is anti-Black because it
1. uses an anti-Black slur in a derogatory way or expresses direct hatred or threats against Blacks.
2. is a derogatory statement generally associating Blacks with crime, drugs, or violence.
3. claims that Blacks are generally lazy or dirty.
4. claims that Blacks are uneducated or stupid.
5. draws stereotypical images of Blacks, such as big lips/butt/breasts.
6. endorses blackface (which has a long history of portraying Blacks in a demeaning way.
7. relates Blacks to monkeys or savages.
8. alleges that black men are predators of white women.
9. alleges that all Blacks grow up without a father.
10. uses "positive" stereotyping, such as all Blacks, are good dancers, fast, and strong.
11. claims that all Blacks share certain opinions/preferences, e.g. love watermelons, or chicken.
12. draws on other anti-Black stereotypes (name in comments).

Anti-Latinx bias is a certain perception of Latinx or Hispanic people, which may be expressed as racial degradation, xenophobia, or hatred toward Latinx. Rhetorical and physical manifestations of anti-Latinx bias are directed toward Latinx or perceived Hispanic individuals and/or their property or toward Hispanic community institutions. Contemporary examples of anti-Latinx bias in public life, the media, schools, the workplace, and in the religious sphere could, taking into account the overall context, include, but are not limited to the examples below.

This tweet is anti-Latinx because it
uses an anti-Latino/Latina slur in a derogatory way or expresses direct hatred or threats against Latinx/Hispanics.
1.  is a derogatory statement generally associating Latinx with crime.
2.  claims that Latinx are illegal or don't pay taxes.
3.  claims that Latinx are lazy.
4.  claims Latinx have too many children.
5.  claims that Latinx steal jobs.
6.  claims that Latinx are uneducated, stupid, or refuse to learn English.
7.  describes Latinas as generally "virginal," "passive," or "dependent on men."
8.  describes Latinas as generally "hot-tempered," "tempestuous," "promiscuous," or exotic seductresses.
9.  describes Latinos as predators of white women.
10. draws on other anti-Latinx stereotypes (name in comments).

Anti-Muslim bias is a certain perception of Muslim people, which may be expressed as hatred toward Muslims. Rhetorical and physical manifestations of anti-Muslim bias are directed toward Muslims or perceived Muslim individuals and/or their property or toward Muslim community institutions and religious facilities. Contemporary examples of anti-Muslim bias in public life, the media, schools, the workplace, and in the religious sphere could, taking into account the overall context, include, but are not limited to the examples below.

This tweet is anti-Muslim because it
1.  uses an anti-Islamic slur in a derogatory way or expresses direct hatred or threats against Muslims or Islam.
2.  is a derogatory statement generally associating Muslims with terrorism and violence.
3.  is a derogatory statement generally associating Islam with terrorism and violence.
4.  claims that Muslims, generally, are dedicated toward the cultural and social destruction of the West.
5.  claims that Muslims, generally, oppress women.
6.  alleges that Muslim men have sex with animals, such as camels or goats.
7.  draws on other anti-Muslim stereotypes (name in comments).

IHRA Working Definition of Antisemitism (detailed description was provided in a link)

Antisemitism is a certain perception of Jews, which may be expressed as hatred toward Jews. Rhetorical and physical manifestations of antisemitism are directed toward Jewish or non-Jewish individuals and/or their property, toward Jewish community institutions and religious facilities.

If you think that this tweet might be antisemitism, select from below.
1. Calling for, aiding, or justifying the killing or harming of Jews in the name of a radical ideology or an extremist view of religion.
2. Making mendacious, dehumanizing, demonizing, or stereotypical allegations about Jews as such or the power of Jews as collective — such as, especially but not exclusively, the myth about a world Jewish conspiracy or of Jews controlling the media, economy, government or other societal institutions.
3. Accusing Jews as a people of being responsible for real or imagined wrongdoing committed by a single Jewish person or group, or even for acts committed by non-Jews.
4. Denying the fact, scope, mechanisms (e.g. gas chambers) or intentionality of the genocide of the Jewish people at the hands of National Socialist Germany and its supporters and accomplices during World War II (the Holocaust).
5. Accusing the Jews as a people, or Israel as a state, of inventing or exaggerating the Holocaust.
6. Accusing Jewish citizens of being more loyal to Israel, or to the alleged priorities of Jews worldwide, than to the interests of their own nations.
7. Denying the Jewish people their right to self-determination, e.g., by claiming that the existence of a State of Israel is a racist endeavor.
8. Applying double standards by requiring of it a behavior not expected or demanded of any other democratic nation.
9. Using the symbols and images associated with classic antisemitism (e.g., claims of Jews killing Jesus or blood libel) to characterize Israel or Israelis.
10. Drawing comparisons of contemporary Israeli policy to that of the Nazis.
11. Holding Jews collectively responsible for actions of the state of Israel.

# Acknowledgements

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. We are grateful that we were able to use Indiana University's Observatory on Social Media (OSoMe) tool and data (Davis et al., 2016).

Contact for more details: Gunther Jikeli, Erna B. Rosenfeld Associate Professor, Institute for the Study of Contemporary Antisemitism | Borns Jewish Studies Program |Germanic Studies | Indiana University, www.isca.indiana.edu