An Unrepresentative Democracy - How Disinformation and Online Abuse Hinder Women of Color Political Candidates in the United States - Codebook

CENTER FOR DEMOCRACY AND TECHNOLOGY

Study Background

The prevalence of disinformation in the U.S. has emerged as a significant problem in recent years, with researchers focusing on the patterns and tactics used by disinformation campaigns. However, much of what we have learned from this research does not adequately examine the specific groups often targeted by disinformation, such as communities of color and women of color in public life. In this research study, we examine the impacts of disinformation specifically on women of color political candidates running for office in comparison to other identity groups.

The main objective of this project is to examine disinformation in an election setting and to contribute meaningfully to the field of research which identifies gendered disinformation as a tool to undermine the political efficacy of women. This will be done using a **content analysis** of public interaction on social media, specifically on the Twitter platform.

Content Analysis and Codebook Use

This project included analysis of public social media content on Twitter. The content analysis was performed *manually* by coders who went through a large-scale dataset of posts on Twitter and classified each tweet on several criteria, as defined in this codebook.

This codebook includes all the categories on which each coder was asked to classify each twitter post, with detailed descriptions on what is and is not included in each category and some examples. The majority of examples are adapted from real tweets. Each coder was asked to read the codebook thoroughly before beginning to code and to use it as a guide at any stage in the coding when the coder was not 100% sure how to classify a particular tweet. Below, we begin with a general guideline on how to classify tweets followed by category-by-category instructions, both of which were given to coders.

Codebook General Guidelines

- **One tweet = one classification:** Please treat each tweet as a single data point. For example, if the tweet includes both text and a photo, you will be asked to identify whether there is any abuse or disinformation towards the candidate in that tweet *as a whole.*

  Example: A tweet mentions that a candidate has spoken publicly against masks and an image presents them wearing a mask. *As a whole*, the author of the tweet suggests that the candidate is attempting to trick her audience or that they are unreliable in what they say.

- **Context of the tweet:** Each tweet will be classified according to the *context* in which the content exists, and not only the content itself. For example, if an individual references someone else's abusive behavior or misinformation sharing and comments on it, but does not *engage in abusive behavior themselves*, that tweet will not be classified as abuse.

- **Tweet subject:** Each tweet will be classified in direct relationship with the *candidate in question.* The candidate in question will be noted on the coding interface for the coders' to see when coding each tweet. For example, if a tweet mentions two candidates, and directs verbal abuse towards one, it will only count as abuse if the *candidate in question* is the one being abused. If the *candidate in question* is the one who is not being abused, the tweet will not be classified as abuse.

- **Long content:** If a tweet includes a long content form, for example, a link to a long article, or a long video, the coder should spend no more than two minutes examining the content and coding it correctly to the best of their ability.

CENTER FOR
DEMOCRACY
& TECHNOLOGY

Codebook Categories

First category: *Is the tweet about the candidate in question?*

Each tweet classification will begin with this question. If the tweet is not directly *about* the candidate in question (as noted in the general guidelines), the coder should answer *no* and will not be required to code any additional category about this tweet. For example, If the candidate is mentioned solely to get them to pay attention to an issue, to ask their opinion about a topic (as a representative, or not), or to comment on an event they will be attending as part of their job, select 'no'. If the tweet *is* about the candidate, the coder should say *yes* and continue to the rest of the classification questions.

All Other Classifications

Below, Each question that is mandatory is marked with an asterisk (*). The ones that are not mandatory do not have an asterisk and should be filled out only if applicable. Each classification category is described below, along with all possible responses.

Name	Description	Example
1 - Sentiment* **(single select)**	What is the stance that is expressed towards the candidate in the tweet? Excludes: sentiment in general (not directly towards the candidate). If it is not clear if the sentiment towards the candidate is intended to be positive or negative, code as neutral.	

Negative	The tweet expresses a stance that opposes the candidate (negative)	@candidate lying about your current job for votes isn't a good look.
Neutral	The tweet does not express a particular positive or negative sentiment towards the candidate (neutral).	If masks work, how did @candidate test positive?
Positive	The tweet expresses a supportive or positive stance towards the candidate	Dear @Twitter, verify @candidate, please ! Her account matters to voters. Thank you!
2 - Narrative* (single-select)	What aspect of the candidate does the tweet narrative focus on? In other words, what is the tweet about? Try to identify the category that best describes the **most significant** narrative of this tweet.	

Identity	The narrative expressed in the tweet is focused on the candidate's perceived or actual personal identity: race, gender, sexual orientation, ethnicity, etc	What happened to my colleague @candidate is despicable. It is vile hatred. This racist verbal assault must be universally condemned.
Ideology	The narrative expressed in the tweet is focused on the candidate's ideology, i.e., ideas that they believe in (politically)	Wow! @candidate sounds like he supports anything that is pro-life.
Policy	The narrative expressed in the tweet is focused on the candidate's political policy and actions taken, including actions they took as part of a government role.	Republican Congressman @candidate removed medical coverage for pre-existing conditions.
Character	The narrative expressed in the tweet is focused on one of the candidate's assumed or actual personality traits.	I enjoyed seeing @candidate come out to talk to the people yesterday. Glad to have them on this mission!

Electability	The narrative expressed in the tweet is focused on whether the candidate is fit or unfit to be elected for the position they are running for. This includes calls for electing this candidate, and support for their campaign. **This category should be selected only if none of the above are evident.** E.g., if the candidate is described as unfit for the job due to their policy, ideology, character or identity, prioritize these categories.	@candidate is simply the wrong man for the job!
2b - Identity Type (multi-select)	If the tweet references the candidate's identity as part of the content, which identity characteristics are in discussion? (select all that apply)	
Disability	The candidate's actual or perceived disability is mentioned or referred to in the tweet (in relation to the candidate).	
Gender	The candidate's gender or perceived gender is mentioned or referred to in the tweet (in relation to the candidate).	There ARE more Women FOR Gun Rights here than for Gun Control. @candidate
Race	The candidate's race or perceived race is mentioned or referred to in the tweet (in relation to the candidate).	
Religion	The candidate's religion or perceived religion is mentioned or referred to in the tweet (in relation to the candidate).	

Sexual Orientation	The candidate's sexual orientation or perceived sexual orientation is mentioned or referred to in the tweet (in relation to the candidate).	
Socio-economic Status	The candidate's current or previous socio-economic status is mentioned or referred to in the tweet.	
Other	Another aspect of the candidate's identity is mentioned or referred to in the tweet.	
3 - Stereotypes (multi-select)	Is any common stereotype related to the candidate's identity expressed in the tweet?	
Women as cold and emotionless	Is the candidate being described as cold and emotionless?	
Women as emotional or hysterical	Is the candidate being described as emotional or hysterical?	
Enemies of the public	Is the candidate described as the public's enemy? (relevant for women, PoC and LGBTQ+ communities)	
Overly sexual	Is the candidate described as overly sexual or promiscuous? (relevant for women, PoC and LGBTQ+ communities)	
Overly concerned with appearance	Is the candidate described as one who is overly concerned with their appearance? (relevant for women, PoC and LGBTQ+ communities)	
Women as radical feminists	Is the candidate being described as a radical feminist? For example, accused of pushing a secret gender equality agenda?	

Women as too motherly	Is the candidate being described as too motherly? (relevant for women and PoC communities)	
Women as unintelligent	Is the candidate depicted as someone who is unintelligent (in comparison to others, or not)	
Violating gender role norms	Is the candidate being described in a way that explicitly violates gender role norms?	
Following gender role norms	Is the candidate being described in a way that explicitly follows their gender role norms?	
PoC as criminals	Is the candidate presented as someone who is involved in criminal activity?	
PoC as idle	Is the candidate presented as particularly idle or "lazy"?	
PoC as unemployed or in blue collar positions	Is the candidate depicted as someone who is unemployed or came from blue collar positions?	
PoC as violent individuals	Is the candidate depicted as someone who is violent towards others?	

PoC as unintelligent	Is the candidate depicted as someone who is unintelligent (in comparison to others, or not)	
Disability or LGBTQ+ identity as illness	Is the candidate's disability or LGBTQ+ identity presented as an illness or as malfunction?	
Person with disability as superhuman	Is the candidate with a disability portrayed as deviant because of 'superhuman feats', or as 'special' because he or she lives a regular life 'in spite of' their disability?	This is amazing – @candidate was able to return to his hobby of playing piano in spite of his recent visual impairments.
Disability as costly	Is the candidate with a disability and their accessibility to society presented as costly to society in general, and to businesses especially?	
Person with disability as "in need"	Is the candidate with a disability presented as disadvantaged and that must look to the state or society for economic support, which is considered a gift, not a right?	
LGBTQ+ as people with more disposable income	Is the candidate described as an individual with more disposable income due to their LGBTQ+ identity?	
LQBTQ+ as hedonistic without responsibilities	Is the candidate described as a hedonistic individual with no responsibilities (sometimes in contrast to their heterosexual counterparts) due to their LGBTQ+ identity?	

Gay men as funny and cheerful	Is the male candidate described as fun and cheerful due to their LGBTQ+ identity?	
Lesbian women wish to look like men	Is the lesbian candidate described with a desire to look like a man due to their LGBTQ+ identity?	
Gay best friend	Is the candidate described as a desired gay best friend for heterosexual women?	@candidate is the perfect fag hag!
Other stereotype	Is the candidate described through another gender / race / disability / LGBTQ+ or other identity stereotype that is not listed above?	
4 - Abuse* (single select)	Is the author of the tweet expressing any kind of abuse towards the candidate? This does not include abuse towards anyone else other than the candidate, or aggressive behavior that is not specifically towards a candidate (e.g., random use of offensive language) Negative sentiment language that is commonly used in politics (such as calling someone a liar or arguing they will not debate you) is not necessarily abuse. Try to focus on instances that are personally offensive, potentially traumatizing, or that may cause an individual to have to avoid or cope with in another way.	
4b - Abuse Type (multi select)	What is the kind of abuse that the author of the tweet is expressing towards the candidate? Select all that apply.	

Demeans or belittles	The tweet demeans or belittles the candidate about their identity, or about other non-professional aspects.	
Direct threat	The tweet expresses a direct physical threat from the person writing the tweet towards the candidate	@candidate you better watch your back… I'm going to kill you and put the video all over the internet. lololol
Indirect threat	The tweet expresses an indirect physical threat towards the candidate	@candidate I hope you enjoy your last moments on earth. You did nothing worthwhile with your life.
Doxing	The tweet shares private information from another platform about the candidate's personal life (content they did not intend to share themselves on Twitter)	
Embarrassing content	The tweet shares embarrassing or cruel content about a candidate	
Impersonation	The tweet uses impersonation to mock the candidate.	
General Offensive language	The tweet uses offensive language, including cursing, rudeness, and other forms of offensive language **directly towards the candidate**. This does not include instances of using offensive language in a non-offensive way (for example: "you are fucking amazing!" (This category not include specifically racist or misogynistic slurs: see categories below)	Speaking as a resident of North Carolina, I can say I'd rather have a senator who sexts than a dumb fuck like @candidate who is reckless about #COVID19.

Sexism or misogyny	The tweet includes insulting or abusive content directed at women based on their gender, including content intended to shame, intimidate or degrade women. It can include profanity, threats, slurs, and insulting epithets.	Ain't nothin worse than a corny chick. – Piss off you disgusting useless fat bitch!
Racism	The tweet includes discriminatory, offensive, or insulting content directed at a person based on their race, including content that aims to attack, harm, belittle, humiliate, or undermine them.	Just be honest and call her what you really want to, an uppity n****r bitch who doesn't know her place.
Homophobia or transphobia	The tweet includes discriminatory, offensive or insulting content directed at an individual based on their sexual orientation, gender identity or gender expression. This includes negative comments towards bisexual, homosexual, and transgender people.	I guess Lesbo's would be into you, I'm sorry not worth my time bitch!
Ethnic or religious slur	The tweet includes discriminatory, offensive, or insulting content directed at a candidate based on their ethnic or religious identities.	Typical muslim full of shit
Promotes or incites violence	The tweet encourages others to harass or physically harm the candidate, or the tweet celebrates or praises acts of violence towards them	
Sexual Assault	The tweet uses sexually abusive language or imagery in turning to, or in talking about the candidate, including unwanted sexual advances.	Would anyone like to fuck @candidate with me? – Shut up or I'll grab you by the pussy
Sexual Content	The tweet uses sexual content related to the candidate as a tool of abuse towards them	Look at that slutty dress @candidate is wearing

Vandalizing	The tweet presents vandalizing related to the candidate	Gotta mess up lame campaigns like @candidate (image of vandalized posters)
Other Form of Harassment	The tweet presents harassment that does not fit into any of the categories above, but still expresses harassment towards the candidate.	

5 - Disinformation * (single select)	Is there dis- or misinformation in this tweet about the candidate, their policy, their party or their campaign? This excludes disinformation or misinformation about another candidate (who is not the topic of the tweet), about a network, social media website, news outlet, etc.
	
	If you are not sure whether this is or is not dis or misinformation, consider the following:
	
	(1) If the information is clearly true, or if it is taken from a reliable source, select 'No'. (not dis- or misinformation)
	(2) If you are not sure if the information is true, or know that it is not true, determine whether the tweet presents dis/misinformation or simply states a personal opinion:
	
	    Ask the question: "If the information in the tweet was to be true, might the information it presents have an impact on the candidacy?"
	
	    (a) If the answer is Yes, select 'Yes' if you know for sure the information presented is not true, or 'unverifiable' if you are not sure.
	    (b) If the answer is No, the tweet is not mis/disinformation – select 'No'.	

| **External Source (open text)** | If you categorized the tweet as "unverifiable", note here if there is any indication of a news source or other source of information included in the tweet (such as another twitter account, a website, a news outlet website etc). This could be in the form of text, in a link, or as part of an image or video. **Add the name of the source domain**: "sourcename.com" | |