Perceiving Politics on TikTok: A User-Centered Approach to Understanding Political Content on TikTok

Jason Greenfield
jg7477@nyu.edu
New York University
New York, New York, USA

Stephanie Wang
stephtw@seas.upenn.edu
University of Pennsylvania
Philadelphia, Pennsylvania, USA

Danaé Metaxa
metaxa@seas.upenn.edu
University of Pennsylvania
Philadelphia, Pennsylvania, USA

Abstract

As TikTok's popularity surges, it faces growing scrutiny over political content on its platform. This study introduces a novel browser-based tool to track users' exposure to and perceptions of political content. We conducted a study with 368 participants, combining an initial survey with data from the custom browser tool. Participants annotated videos in real-time logged-in sessions, indicating whether they perceived each of 40 sequential videos in their personalized TikTok feeds as political, and providing brief justifications. Using expert and LLM coding alongside natural language processing techniques, we identify and characterize political topics users encounter. Despite TikTok's growing role as a news source, we observe few official channels for news, and contrary to public concerns, users do not report high volumes of politically extreme content. We conclude with implications for user-centered studies of political content online and directions for future research.

Keywords

TikTok, social media, political content.

ACM Reference Format:
Jason Greenfield, Stephanie Wang, and Danaé Metaxa. 2025. Perceiving Politics on TikTok: A User-Centered Approach to Understanding Political Content on TikTok. In *Proceedings of Conference on Human Factors in Computing Systems (CHI '25).* ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

With over a billion and half users worldwide, TikTok is surging in global popularity, capturing vast sectors of younger internet users and other demographic groups [8]. While a third of U.S. adults have reported using TikTok, we know relatively little about what political content exists on the platform, how much of it users meaningfully engage with, or how users perceive this content [13]. Meanwhile, U.S. lawmakers have expressed concerns that TikTok poses a national security threat due to ownership by a Chinese parent company, ByteDance [23]. These allegations go so far as to suggest that TikTok could be knowingly pumping malicious content through its platform to U.S. audiences, whether Chinese propaganda or politically extreme and polarizing content [7]. Even though TikTok has established trust and safety initiatives, like its Transparency and Accountability Center [26], there is still limited independent research on what political content exists on TikTok.

We present a novel study that takes a user-centered view of what political content people see on TikTok. We conduct a survey with 368 TikTok users based in the U.S., asking questions about participants' interest in politics and how they engage with political content on TikTok. We combine these survey responses with data collected by our custom browser extension, which collects a portion of participants' real TikTok feeds, along with their annotations of whether each video they view in their personalized feed is political, in their own perception. We use this data to ask: 1) *What kinds* of content on the platform do TikTok users perceive to be political? 2) *What specific* TikTok videos do users perceive to be political, and what are their characteristics? To answer these questions, we hand-code participant descriptions of the political content they encounter on TikTok and develop a taxonomy of political topics. We apply this taxonomy to code participant annotations of political videos encountered on their feed, as well as the videos' transcripts.

2 Background, Motivations and Related Work

The rise of TikTok as a platform for political engagement and information-sharing is a growing and important area of study, particularly given its popularity among younger demographic groups and its use as a source of news media. Recent studies by Pew Research Center have identified that 4 in 10 young adults in the U.S. regularly get news from TikTok[18], and that TikTok (52%) has surpassed Facebook (48%) and is approaching Twitter/X (59%) in the amount of users reporting they regularly get news on the platform [18]. One recent study comparing TikTok to YouTube found that TikTok encourages higher rates of content creation among its user base, and that view counts on the platform are determined more by virality than creator popularity [15]. This emphasis on virality could enhance the reach of political content across younger age groups, especially if they are embedded in memes or entertainment-based formats [3, 30], given that TikTok has become a central platform for youth political engagement and activism [17, 19].

Empirical research about news and politics on the platform remains rare. In one landmark study, Medina Serrano et al. [21] scraped a large dataset of U.S. partisan videos marked with political hashtags, finding that political communication on TikTok is interactive and popular with younger age groups. Internationally, Berdón-Prieto et al. [2] collected videos from Spanish political parties, finding them posting increasingly polarizing messages, but

CHI '25, April 26–May 1, 2025, Yokohama, Japan

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

another content analysis study of several populist right-wing parties' posts found that extreme content might not drive engagement as much as humor and entertainment on the platform [11].

Prior work about political social media content, whether on TikTok or other platforms, largely relies on a top-down determination of what constitutes political content [12, 14], or explicitly focuses on news content as a proxy for political content or content of political import (e.g., [22, 25, 28]). While there has been research trying to capture popular conceptions of politics [29] or democracy [4], limited research has applied this question to social media. One area in which differing opinions about political content online have become salient is political advertising. This includes cases like Twitter/X's 2019 ban on political advertising [10] (a decision since reversed [6]), or national governments' passage of regulation on political advertisements [9]. Demonstrating the subjectivity and difficulty in effectively categorizing political content, such cases have sparked interrogations from the popular press [5], think tanks [16, 20], and academics [9, 24]. All these bodies of literature illustrate that TikTok itself and user perceptions of political content on the platform are relatively understudied. In this work, we aim to collect empirical data to grow the field's user-centered understanding of both.

3 Data

Our data set comprises a combination of survey responses, videos scraped from our respondents' TikTok feeds via our custom browser extension, and participant annotations of those scraped videos. Our data collection ran from July 2023 to September 2023 with 368 participants recruited through Prolific, a common research participation site. The respondents included in our final dataset are similar to U.S. demographics by sex (53.35% Male, 42.74% Female, 3.63% Not provided, 0.28% preferred not to say) and race (65.64% White, 28.21% People of Color, 6.15% Not provided). We actively recruited for an ideologically balanced group of participants (48.89% Liberal, 41.06% Conservative, 10.06% Neutral).[1] Our respondents first answered a series of survey questions related to their demographics, interest in politics, and use of the TikTok platform. After completing our survey, eligible users (U.S. adults who had used the platform for at least a month and were willing to use Google Chrome to complete the study) were asked to download our custom Chrome browser extension. Upon doing so, they were given instructions to log in to their TikTok accounts in the same browser and scroll through their TikTok feed. For each video, our extension embedded a radio button and open-response box to the left of each video (Figure 1); these asked participants whether or not each video was political (in their own opinion), and to briefly describe why or why not. Participants were required to annotate at least 40 videos (regardless of the proportion they marked as political).

After data cleaning, the *respondent-annotated* dataset comprised 16,332 videos from 358 participants, with 2,171 flagged by respondents as political (13.3%). We downloaded the .mp4 video files for

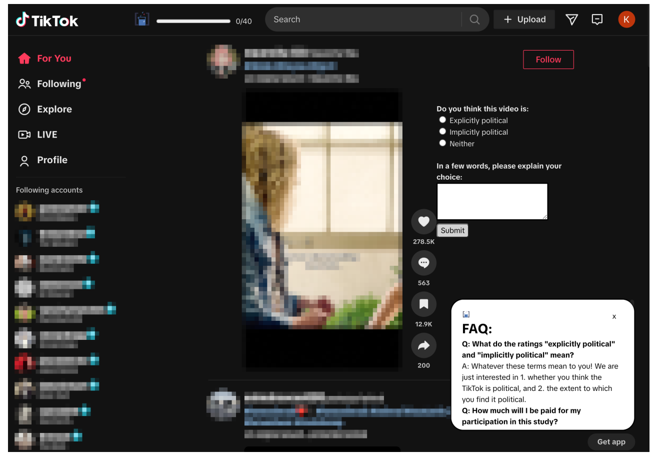

Figure 1: Screenshot of the web tool running on Tiktok. Users can rate the video with one of the three radio buttons, explain their choice, and submit the answer. Additionally, they can see how many video they rated with the progress bar to the right of the Tiktok logo and scroll through the Frequently Asked Questions section on the bottom right of the page.

all videos annotated by users as political that were available (n = 1,616).[2] We further processed the videos by first extracting the audio from the video, then classifying the audio using Audio Spectrogram Transformer (AST)[3] into classes to identify audio with spoken language. Audio classified as containing speech were then transcribed using OpenAI's Whisper model (medium size).[4] The final *respondent-political-transcribed* dataset contains 1,115 transcribed videos (across 19 different languages) of the 2,171 marked by respondents as political.

4 Coding Survey Responses and User Annotations of TikToks

To understand what kinds of content on TikTok participants perceive to be political, we code respondent data for the presence of political themes and topics. We develop a codebook of 30 political codes and then apply it to three different datasets: (1) survey responses, (2) user annotations from the *respondent-political-transcribed* dataset, and (3) video transcripts from the same.

We start by qualitatively coding user responses (n = 368) to our survey question concerning perceptions of political content on TikTok: "In two sentences or more, how would you describe the kinds of TikTok content you consider to be political, and what makes it political to you?" To develop the codebook, two of the authors followed an iterative inductive coding process using the survey responses. Because the dataset was relatively small, the authors were able to each code it in its entirety during each round of coding. In each round, they independently read and annotated survey response data to identify new codes and refine the definitions of existing codes present in the data. After each round of coding,

[1]We note that this sample is certainly not representative of the base of all TikTok users, even within the U.S.; for example, 58% of U.S. teens report using TikTok on a daily basis [1], whereas our participants are restricted to legal adults. Moreover, we (and perhaps even TikTok itself) lack the demographic data about the full user base needed to gather a representative sample. Instead of representativeness, then, we aimed to recruit an ideologically, racially, and gender diverse pool.

[2]Some videos were taken down from the platform between the time of the survey and our data collection, or were unavailable for downloading for some other reason.
[3]https://huggingface.co/MIT/ast-finetuned-audioset-10-10-0.4593
[4]https://github.com/openai/whisper

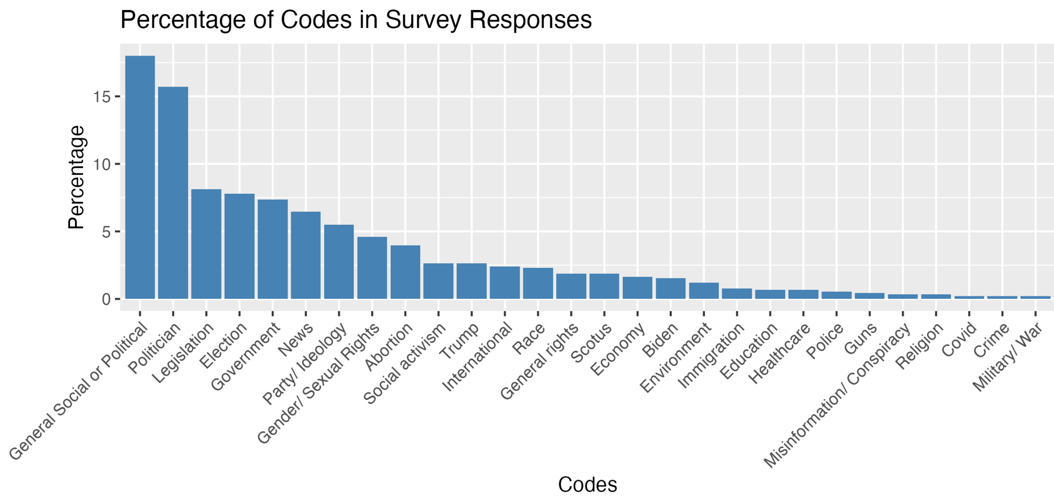

Figure 2: Frequency of all codes in participant descriptions of political content on TikTok. The highest frequency codes suggest that participants use general language to describe political content instead of references to specific policy issues.

the two coders met to discuss trends in the data and refine the codebook. The codebook was finalized after four rounds, when saturation had been reached (i.e., no new codes were being added) and inter-rater reliability was high (Cohen's $\kappa \geq 0.8$ for each code). The full list of codes and inter-rater reliability measures is included in Appendix A.

After the codebook was developed by coding the survey responses, it was applied in combination with a Large Language Model (LLM) to code the user annotations and video transcripts from the *respondent-political-transcribed* dataset. We started by hand-coding a random sample (n = 150) of annotation-video pairs from the *respondent-political-transcribed* dataset, coding both the user annotations and the videos themselves. During this process we augmented our survey-based codebook in two cases where our existing codes did not suffice, adding new codes PRISON and HOUSING. We then compared each set of our expert-generated codes to codes generated by ChatGPT for user annotations and video transcripts using the following prompt, where CODEBOOK is a dictionary mapping of the code to its definition (Appendix A), and TEXT is either the user free text response or video transcript:

> Given the following predefined codes for political content: {CODEBOOK} Classify the following text into one or more of these categories: "{TEXT}" Respond with ONLY the category name (e.g., TRUMP, CDP, MC, IMG).

After prompt engineering and minor refinements to the codebook, we achieved substantial inter-rater reliability between the expert and machine-coded values for 22 codes (Cohen's $\kappa \geq 0.75$ for each code on labels of both user annotations and video transcripts).

4.1 Validating User Annotations of TikToks

While coding *respondent-political-transcribed* for political topics, we noted the presence of videos that were not political which we code as NONE. We remove videos annotated by users as political that are not actually political from the *respondent-political-transcribed* dataset by comparing the codes in the short-form descriptions of why a user found a video to be political with the "ground truth" codes generated from that video's text transcript. If both the code of the user's short-form description and the code of the video transcript are NONE, meaning that none of the 30 political topics defined in Section 4 are present in either pieces of content, we remove the video from the dataset. We identify 138 such annotation-video pairs, and after removing them from *respondent-political-transcribed* we are left with the *validated-perceived-political* dataset (n = 997) which we use in our subsequent analysis.

5 Initial Results

We examine the relative presence of the codes developed in Section 4 in each of the following: (1) survey responses, (2) user annotations, and (3) video transcripts. As a reminder, the survey responses consist of free-text descriptions in which participants describe the kinds of content on TikTok they consider political, the user annotations contain direct assessments of specific TikTok videos that participants encountered and marked as political, and the video transcripts are textual transcriptions of TikTok videos classified as political based on our coding scheme. In this section, we describe the frequency of codes in each dataset; figures showing the full set of frequencies for each are available in Appendix B. Finally, we further characterize political TikTok videos by analyzing the emotional and sentiment-based framing that co-occurs with each code.

In analyzing survey responses, we find that respondents primarily reference established political entities rather than specific policy issues. Among the most frequent codes mentioned by participants was General Social or Political, which refers to mentions of social

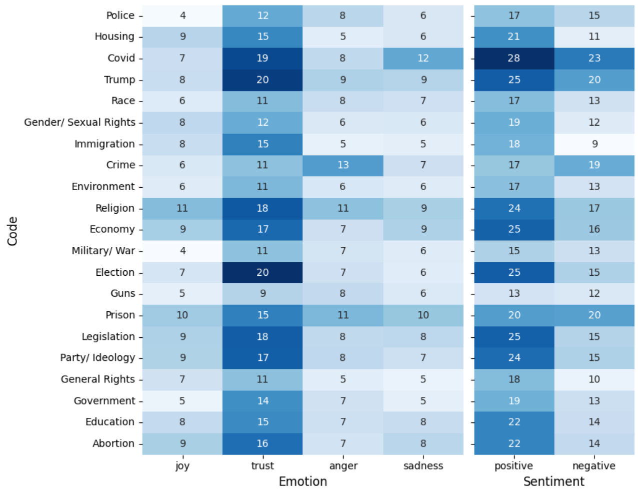

Figure 3: Heatmap displaying the co-occurrence of codes in political video transcripts with values for emotion and sentiment. Political TikToks co-occur with high levels of trust, and generally more positive than negative sentiment.

and political topics in vague or general terms; for example, participants answering, *"Anything talking about politics/social issues"* or *"If it deals with political messages or themes"*. Other high-frequency codes include Government, Elections, Legislation, and Politicians. Notably, more specific issue-related codes, such as Economy, Abortion, and Immigration, are mentioned far less frequently. The first mention of a specific policy issue does not appear until the eleventh most frequent code, Gender and Sexual Identity Rights, which often co-occurs with other codes and is frequently referenced in relation to political figures. For example, one participant described, *"I see TikToks about Trump and trans-rights issues. These are political because Trump is a political figure and trans rights are a political talking point."* Interestingly, codes related to misinformation (Misinformation/Conspiracy) appear in less than 2% of survey responses.

Turning to user annotations, we find that the most frequently applied codes are issue-specific. The most common code is Race, as seen in annotations such as *"This is about systemic racism."*, followed by Gender and Sexual Identity Rights, *"LGBT person being discriminated against at work"*. Next are Police , *"police brutality"*; Crime, *"deals with a case of assault"*; and Economy, *"This video speaks of the financial strain of inflation in America on young people."*

Finally, we compare codes to the emotion and sentiment present in videos in order to better characterize the communication style of TikTok videos perceived as political. We use the EmoLex Python package [5] to generate scores for four emotional values in addition to the overall positive and negative sentiment for each video in the *validated-perceived-political* dataset based on its text transcript. Figure 3 shows the co-occurrence of transcript-derived codes and scores for emotion and sentiment. We observe that political TikTok videos tend to be more positive than negative, with high values for trust and low values for anger and sadness. We also observe that Covid, Prison, Trump, and Crime have the highest scores for negative sentiment and all but Crime have equal or greater values for positive sentiment.

6 Discussion and Conclusion

Our initial analysis identifies over 30 political topics participants perceive to be political on TikTok. When asked in our survey to describe the political content they see, participants used general language and were most likely to mention governmental and legal entities (e.g., Politician, Legislation, Election, Government). In contrast, when faced with specific videos in their feeds, these topics were much less common compared with issue-specific language (e.g., Race, Gender/ Sexual Rights, Police, Economy, Crime). This discrepancy suggests a saliency bias towards formal themes compared to specific issues when participants are tasked with describing this content in the abstract. We recommend future user-centered work on political social media content instead use concrete examples and collect data sets of actual social media content.

Additionally, we show that the majority of political content is associated with positive sentiment, which aligns with survey data showing that 84% of U.S. adult TikTok users encounter humorous posts referencing current events, and 80% see content expressing opinions about them [27]. On the other hand, the topics exhibiting the most intense displays of sentiment, measured by the combined levels of positive and negative sentiment, are COVID and TRUMP. This suggests that while political content on TikTok often leans positive, polarizing topics are still portrayed with strong sentiment. In future work, we plan to conduct a deeper content analysis of the videos that users flag as political, and also to examine whether any differences in the frequency or emotional affect of codes in perceived-political videos can be attributed to user demographics or survey response factors.

Given TikTok's rising popularity as a news source for U.S. adults, we aimed to identify NEWS content in our data but encountered challenges in reliably coding it. Our LLM-based pipeline yielded low inter-rater reliability scores for the NEWS code in both user annotations and text transcripts, leading us to exclude it from the analysis. This likely stems from the complexities of coding multimodal data. For example, a video featuring a clip from The Rachel Maddow Show was manually coded as NEWS, but the machine coding missed this because it relied solely on the text transcript, which lacked key visual cues, such as the news show host. In manual coding, we also examined whether video posters were linked to official news media or unofficial accounts like influencers. This initial analysis found that few NEWS content posters were from official news accounts, even when using full clips from news broadcasts. We hypothesize that this is evidence that traditional news media accounts have yet to fully breached the TikTok platform despite the high number of Americans reporting they regularly get news from the platform [18]. As news content is commonly used as a proxy for political content more generally in social media research, our finding suggests such approaches may be challenging to implement and miss the bigger picture if applied to TikTok.

6.1 Limitations

First, the scope is limited by a small sample size and a U.S.-focused population. While we aimed for demographic diversity, TikTok's

[5] https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

global reach means users from different regions may have varying perceptions of political content. Future research could recruit geographically diverse users to compare these perceptions. Second, we did not fully explore the impact of demographic variables on political content perceptions. Future studies could examine how these perceptions vary across demographic groups and relate to survey-reported values. Additionally, our analyses currently focus only on the textual elements of TikTok videos. In future work, we plan to incorporate the full multimodal nature of the data, including visual and auditory components. Another limitation is the reliance on self-report data. We think it is important to take a human-centered focus when describing perceptions of politics, but there may be categories of content (like misinformation, for example) that participants may not have described as such. In future work we plan to develop a classifier to identify political videos participants may have missed during annotation, and develop a dataset of ground-truth political videos. This dataset can be used to assess the accuracy and completeness of participant annotations. By comparing self-reported data with classifier-generated labels, we can identify potential gaps in perception and better understand how different groups interpret political content. We can also gain a clearer understanding of how demographic factors influence user exposure to various types of political content.

References

[1] Monica Anderson, Michelle Faverio, and Jeffrey Gottfried. 2023. Teens, social media and technology 2023. *Pew Research Center* (2023).

[2] Pablo Berdón-Prieto, Jacobo Herrero-Izquierdo, and Itziar Reguero-Sanz. 2023. Political polarization and politainment: Methodology for analyzing crypto hate speech on TikTok. *Profesional de la información* 32, 6 (2023).

[3] Yanni Brown, Barbara Pini, and Adele Pavlidis. 2024. Affective design and memetic qualities: Generating affect and political engagement through bushfire TikToks. *Journal of sociology* 60, 1 (2024), 121–137.

[4] Damarys Canache. 2012. Citizens' conceptualizations of democracy: Structural complexity, substantive content, and political significance. *Comparative Political Studies* 45, 9 (2012), 1132–1158.

[5] Kate Conger. 2019. What Ads Are Political? Twitter Struggles With a Definition. https://www.nytimes.com/2019/11/15/technology/twitter-political-ad-policy.html Accessed: 2024-10-30.

[6] Sheila Dang. 2023. More Americans – especially young adults – are regularly getting news on TikTok. https://www.reuters.com/business/media-telecom/twitter-expand-permitted-political-advertising-2023-01-03/ Accessed: 2024-10-30.

[7] Ken Dilanian. 2024. TikTok says it's not spreading Chinese propaganda. The U.S. says there's a real risk. What's the truth? *NBC News* (16 September 2024). https://www.nbcnews.com/investigations/tiktok-says-not-spreading-chinese-propaganda-us-says-real-risk-rcna171201

[8] Stacy Jo Dixon. 2024. https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

[9] Katharine Dommett and Junyan Zhu. 2023. What is an online political advert? An interrogation of conceptual challenges in the formation of digital policy response. *Policy & Internet* 15, 4 (2023), 713–730.

[10] Jack Dorsey. 2019. We've made the decision to stop all political advertising on Twitter. https://x.com/jack/status/1189634360472829952 Accessed: 2024-10-30.

[11] Juan Manuel González-Aguilar, Francisco Segado-Boj, and Mykola Makhortykh. 2023. Populist right parties on TikTok: Spectacularization, personalization, and hate speech. *Media and communication* 11, 2 (2023), 232–240.

[12] Sandra González-Bailón, David Lazer, Pablo Barberá, Meiqing Zhang, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Deen Freelon, Matthew Gentzkow, Andrew M. Guess, Shanto Iyengar, Young Mie Kim, Neil Malhotra, Devra Moehler, Brendan Nyhan, Jennifer Pan, Carlos Velasco Rivera, Jaime Settle, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Chad Kiewiet de Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, and Joshua A. Tucker. 2023. Asymmetric ideological segregation in exposure to political news on Facebook. *Science* 381, 6656 (July 2023), 392–398. https://doi.org/10.1126/science.ade7138 Publisher: American Association for the Advancement of Science.

[13] Jeffrey Gottfried. 2024. Americans' Social Media Use. https://www.pewresearch.org/internet/2024/01/31/americans-social-media-use/ Accessed: 2024-10-29.

[14] Andrew M. Guess, Neil Malhotra, Jennifer Pan, Pablo Barberá, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Drew Dimmery, Deen Freelon, Matthew Gentzkow, Sandra González-Bailón, Edward Kennedy, Young Mie Kim, David Lazer, Devra Moehler, Brendan Nyhan, Carlos Velasco Rivera, Jaime Settle, Daniel Robert Thomas, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Beixian Xiong, Chad Kiewiet de Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, and Joshua A. Tucker. 2023. How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science* 381, 6656 (July 2023), 398–404. https://doi.org/10.1126/science.abp9364 Publisher: American Association for the Advancement of Science.

[15] Benjamin Guinaudeau, Kevin Munger, and Fabio Votta. 2022. Fifteen seconds of fame: TikTok and the supply side of social video. *Computational Communication Research* 4, 2 (2022), 463–485.

[16] Julian Jaursch. 2020. Defining Online Political Advertising. *interface* (2020). https://www.interface-eu.org/storage/archive/files/snv_definingpoliticalads.pdf

[17] Kiana Karimi and Richard Fox. 2023. Scrolling, Simping, and Mobilizing: TikTok's influence over Generation Z's Political Behavior. *The Journal of Social Media in Society* 12, 1 (2023), 181–208.

[18] Rebecca Leppert and Katerina Eva Matsa. 2024. More Americans – especially young adults – are regularly getting news on TikTok. https://www.pewresearch.org/short-reads/2024/09/17/more-americans-regularly-get-news-on-tiktok-especially-young-adults/ Accessed: 2024-10-27.

[19] Ioana Literat and Neta Kligler-Vilenchik. 2023. TikTok as a key platform for youth political expression: Reflecting on the opportunities and stakes involved. *Social Media+ Society* 9, 1 (2023), 20563051231157595.

[20] Tomas Apodaca Maria Puertas. 2024. Instagram Creators: Check If Your Posts Are "Political". https://themarkup.org/automated-censorship/2024/03/25/instagram-creators-check-if-your-posts-are-political

[21] Juan Carlos Medina Serrano, Orestis Papakyriakopoulos, and Simon Hegelich. 2020. Dancing to the Partisan Beat: A First Analysis of Political Communication on TikTok. In *Proceedings of the 12th ACM Conference on Web Science* (Southampton, United Kingdom) *(WebSci '20)*. Association for Computing Machinery, New York, NY, USA, 257–266. https://doi.org/10.1145/3394231.3397916

[22] Brendan Nyhan, Jaime Settle, Emily Thorson, Magdalena Wojcieszak, Pablo Barberá, Annie Y Chen, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Drew Dimmery, et al. 2023. Like-minded sources on Facebook are prevalent but not polarizing. *Nature* 620, 7972 (2023), 137–144.

[23] U.S. House of Representatives. 118th Congress 2nd Session. 2024. *H.R. 7521, Protecting Americans from Foreign Adversary Controlled Applications Act.*

[24] Orestis Papakyriakopoulos, Christelle Tessono, Arvind Narayanan, and Mihir Kshirsagar. 2022. How Algorithms Shape the Distribution of Political Advertising: Case Studies of Facebook, Google, and TikTok. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (Oxford, United Kingdom) *(AIES '22)*. Association for Computing Machinery, New York, NY, USA, 532–546. https://doi.org/10.1145/3514094.3534166

[25] Ronald E. Robertson, Jon Green, Damian J. Ruck, Katherine Ognyanova, Christo Wilson, and David Lazer. 2023. Users choose to engage with more partisan news than they are exposed to on Google Search. *Nature* 618, 7964 (June 2023), 342–348. https://doi.org/10.1038/s41586-023-06078-5

[26] TikTok. 2024. Transparency Center. https://www.tiktok.com/transparency/en-us/ Accessed: 2024-10-29.

[27] Luxuan Wang, Galen Stocking, Samuel Bestvater, and Regina Widjaya. 2025. Americans' experiences with news on TikTok: A closer look | Pew Research Center. https://www.pewresearch.org/short-reads/2025/01/17/a-closer-look-at-americans-experiences-with-news-on-tiktok/

[28] Stephanie Wang, Shengchun Huang, Alvin Zhou, and Danaë Metaxa. 2024. Lower Quantity, Higher Quality: Auditing News Content and User Perceptions on Twitter/X Algorithmic versus Chronological Timelines. 8, CSCW2, Article 507 (Nov. 2024), 25 pages. https://doi.org/10.1145/3687046

[29] Mark E Warren. 1999. What is political? *Journal of theoretical politics* 11, 2 (1999), 207–231.

[30] Jing Zeng and Crystal Abidin. 2021. '#OkBoomer, time to meet the Zoomers': studying the memefication of intergenerational politics on TikTok. *Information, Communication & Society* 24 (08 2021), 1–23. https://doi.org/10.1080/1369118X.2021.1961007

A Codebook: Participant Descriptions of Political Content on TikTok

The following is the codebook used to label participant descriptions of political content on TikTok. The table includes the abbreviated name of the code, full name of the code, inter-coder reliability measure (Cohen's κ) on initial survey responses, and full definition.

Abbreviated Code	Full Code, Cohen's κ	Definition
GR	General Rights, 1.000	Includes any mention of labor rights or rights for demographics which are not included in another code (i.e. trans-rights would be GSR). Anything about working conditions is also GR (not ECON).
RACE	Race/Ethnicity/Minority, 1.000	Includes any mention of rights or issues related to race. Also includes ethnicity or general mentions of minority groups.
GUNS	Guns, 1.000	Includes any mention of gun control and gun rights. Any response with a mention of guns is coded as GUNS.
ECON	Economy, 0.921	Includes any mention of something having to do with the economy. Mentions of taxes, minimum wage, inflation, labor, product boycotts are all sufficient. Note, mentions of taxes (L), student loans (EDU), and labor strikes (SA) are also coded as ECON.
A	Abortion, 0.983	Includes any mention of abortion.
GSR	Gender / Sexual Identity Rights, 0.986	Includes any mention of gender, sexuality, and LGBT issues.
POLC	Police, 1.000	Includes any mention of police or law enforcement.
CRIME	Crime, 1.000	Includes any mention of crime, including generic references to criminal justice. Mentions of police should be coded as police, and not crime in addition.
MW	Military or War, 1.000	Includes any mention of the military or war.
R	Religion, 1.000	Includes any mention of religion.
EDU	Education, 1.000	Includes any mention of schools, education policy, or student loans.
ENV	Environment, 1.000	Includes any mention of the environment or sustainability.
HC	Healthcare, 1.000	Includes any mention of hospitals, or healthcare.
IMG	Immigration, 1.000	Includes any mention of immigration.
HOUSING	Housing, –	Mentions of house prices, rent, gentrification. Do not mark these as ECON.
PRISON	Prison, –	Mentions of prisons, incarceration.
COVID	COVID-19, 1.000	Includes any mention of COVID-19.
EF	Election Fraud, 1.000	Includes any mention of election fraud.
MC	Misinformation or Conspiracy, 1.000	Includes any mention of misinformation / conspiracies.
E	Election, 0.957	Includes any mention of election-related content. Mentioning "politicians post more of these videos around elections" does not count as this category since the response does not use elections to define political content. Mentioning a "political candidate" should be coded as both election (E) and politician (PO). References to political debates are also included in E.
SA	Social Activism, 1.000	Includes any mention of activism, and references to collective action by individuals (strikes, boycotts, marches, protests, etc.).
NEWS	News, 0.980	Includes any references to "news" itself or mentions of "current events".
GSP	General Social or Political, 0.767	Includes any mentions or generic reference to something such as a "political issue" or "hot topic" or "social issues" or "politics in general." Or, if the participant implies that political content is a general category or wide range of things (i.e., no binary bucket for what is political content). This is a catch all for any generically phrased reference in the response about political or social issues without explicitly mentioning one (though it can occur in the same response as explicit mentions to social or political issues). Also mentions of "current political issues" and NOT "current political events" or similarly general references to political topics are treated as GSP and not NEWS. When there's references to a few specific things as examples or a general or more broader category it's GSP.

Continued on next page

Table 1: (continued)

Abbreviated Code	Full Code, Cohen's κ	Definition
INTL	International, 0.974	Includes any mentions of other countries outside of the US. Also includes generic language specifying countries in general (e.g., "Any content that is focused on large cultural issue (e.g. racism, rights, etc.) that is related to a specific country I consider political.").
PO	Politician, 0.928	Includes any mention referencing a politician. Mentioning a "political candidate" should be coded as both election and politician. Does not include Trump/Biden since they are separate codes.
TRUMP	Trump, 1.000	Includes any keyword mention of TRUMP. Should not be coded as PO, only code as TRUMP.
BIDEN	Biden, 1.000	Includes any keyword mention of BIDEN. Should not be coded as PO, only code as BIDEN.
PARI	Party or Ideology, 0.977	Includes any keyword mention of specific party entities (i.e., democrats, republicans), or specific party entities (i.e., DNC, RNC), or mentions of specific political ideologies (i.e., liberal, conservative). (Note: 'videos persuading you of liberal values' would be coded as both PARI and O; PARI for mention of 'liberal values', and O for the persuasion of an opinion). Generic mentions of something like a 'political stance', simply "political party" is not sufficient. Mentions of "left", "right", "liberal", "conservative" are all sufficient.
G	Government, 0.943	Includes any mention of the government in general or institutions. Does not include POLC or SCOTUS.
SCOTUS	Supreme Court, 0.966	Includes any mention of the Supreme Court. Note that SCOTUS entries should not be coded as G unless there is a specific mention of government or government institutions besides the Supreme Court.
L	Legislation, Laws, or Bills, 0.965	Includes any mention of legislation, laws, bills, and generic mentions of 'policy' that the government enacts/enforces. If taxes are mentioned as a law/policy, it is coded as both ECON and L.
CDP	Controversial, Divisive, Polarizing, 1.000	Includes any references to content that is C/D/P. The lower bound for POLARIZING is any reference to pushing people into two party positions (i.e., 'all political content pushed you to support one belief or another'), and anything that references an "us versus them" mentality. Any references to two parties arguing with each other over a political topic should also be coded as CDP. The lower bound for DIVISIVE is someone being mad or hating another (i.e., 'one politician arguing with another' or 'one politician pointing out what's wrong with another's platform' are not sufficient). Mentions of screaming or extreme behavior or saying something is "extremely political" is CDP. Simple mentions of extreme subjects like Hitler, Nazis, or White Supremacists do not get coded as CDP unless these groups are doing something (i.e., 'white supremacists arguing with a protester' should be coded as CDP but a simple mention of 'white supremacy' should not be coded as CDP). Mentions of propaganda are not coded as CDP. Name calling in itself, or being opposed to someone else should not be coded as CDP—these actions are not controversial.
NONE	None, 0.989	No mention of any definition about politics at all (i.e., 'funny animal tiktoks and wholesome couple tiktoks') should be coded as NONE.

B Distribution of Codes in User Data

This appendix includes full details on the frequency of codes present in: (1) user annotations, and (2) video transcripts.

B.1 What Specific TikTok Videos do Users Perceive to be Political?

Figure 4 shows the frequency of all codes in user annotations describing why they thought a particular video is political.

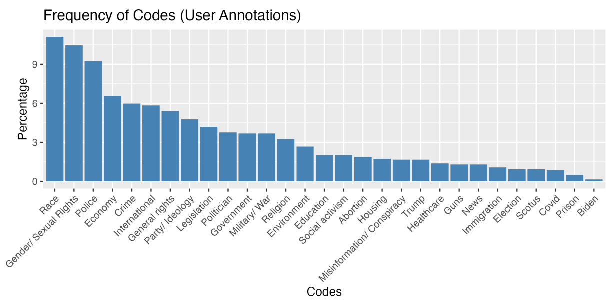

Figure 4: Frequency of all codes (N = 1202) in user annotations describing why they thought a particular video is political.

B.2 What Specific TikTok Videos are Political?

Figure 5 shows the frequency of all codes in text transcripts of political videos.

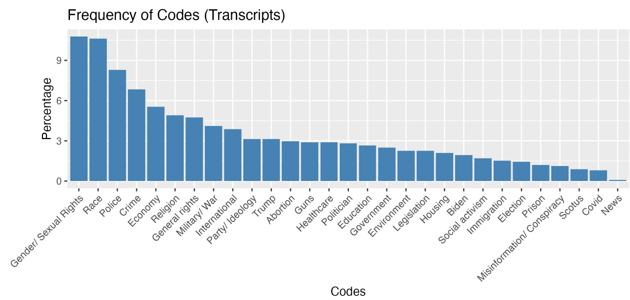

Figure 5: Frequency of all political codes (N = 1088) in video transcripts.

Received 27 February 2025