The CCP's Digital Charm Offensive: How TikTok's Search Algorithm and Pro-China Influence Networks Indoctrinate GenZ Users in the United States

PRESENTED BY

Joel Finkelstein, Ph.D.
Chief Science Officer and Director, Network Contagion Research Institute; Senior Research Fellow, Miller Center for Community Protection and Resilience, Rutgers University

Lee Jussim, Ph.D.
Former Chair, Distinguished Professor,
Department of Psychology, Rutgers University, Chair of Anthropology

John Farmer
Former New Jersey State Attorney General and Senior Counsel, 9/11 Commission;
Director, Miller Center for Community Protection and Resilience, Rutgers University

Alex Goldenberg
Lead Intelligence Analyst, Network Contagion Research Institute; Research Fellow, Miller Center for Community Protection and Resilience, Rutgers University

Jacob Zucker
Senior Intelligence Analyst, Network Contagion Research Institute

Danit Finkelstein
Ph.D. Student, Social Psychology, Rutgers University

Sonia Yanovsky
Ph.D. Student, Social Psychology, Rutgers University

Analysts
Anisha Jagdeep
Ankita Jagdeep
Collin Vasko
Mikey Sohn

TABLE OF CONTENTS

EXECUTIVE SUMMARY

In December 2023, the Network Contagion Research Institute (NCRI) published an initial report that found a strong possibility of content on TikTok being either amplified or suppressed based on its alignment with the interests of the Chinese Government. The report concluded with a warning: “should future research find that TikTok users exhibit attitudes and assessments of world events aligned with these information distortions, democracies will need to consider appropriate countermeasures to protect information integrity and mitigate potential real-world impacts.” These findings underscored the urgency of investigating the specific mechanisms and broader implications of potential algorithmic manipulations.

While our initial study shed light on the potential for content manipulation on TikTok, it did not explore whether specific moderation algorithms or practices are used to suppress topics sensitive to the Chinese Communist Party (CCP). Moreover, it did not compare the nature and impact of pro- and anti-CCP content across other major social media platforms like Instagram and YouTube.

This current report addresses these gaps by providing a more comprehensive analysis of TikTok's moderation practices, examining the nature and prevalence of CCP-sensitive content, and evaluating how different platforms handle such content. Furthermore, this study explores the association between social media platform usage and pro-CCP attitudes among users. By combining user journey data, survey research, and cross-platform comparisons, we uncover the extent to which TikTok and other platforms may influence user perceptions and behaviors in favor of the CCP.

The publication of this research precedes the delivery of oral arguments for *TikTok Inc. and ByteDance Ltd. v. Merrick B. Garland*, a federal court case that has subjected the TikTok algorithm to intensive new scrutiny. As detailed in the DOJ’s response to TikTok’s lawsuit, “covert content manipulation” is an area of legal interest for the DOJ, not to mention its relevance to the broader U.S. intelligence and national security community. That said, within the DOJ response it is noted that definitively *proving* covert content manipulation by the Chinese government remains a challenge. As FBI Director Wray testified to in multiple hearings, “[t]his sort of manipulation of the algorithm would be difficult to detect…"[1] This research gap sets the stage for our findings, which, while not definitive proof of state orchestration, present compelling and strong circumstantial evidence of TikTok’s covert content manipulation.

In response to the DOJ's acknowledgment of challenges in proving manipulation by the Chinese government, our findings present the following circumstantial evidence:

- **Algorithmic Bias:** TikTok's algorithms consistently amplify pro-CCP content and suppress anti-CCP narratives.
- **Content Origination:** Much of the pro-CCP content originates from state-linked entities, including media outlets and influencers.

[1] U.S. Department of Justice. (2024, July 26). *Response to TikTok lawsuit,* pp. 38. DocumentCloud. https://www.documentcloud.org/documents/25017864-72624-doj-response-to-tiktok-lawsuit

- **Psychological Impact:** Survey data shows significant shifts in user attitudes towards China, especially among heavy TikTok users, indicating successful indoctrination.

These points collectively indicate a systematic manipulation of information, suggesting that propaganda produced by state actors and orchestrated through assets owned or influenced by them shapes user perceptions at a massive scale.

BLUF (Bottom Line Up Front)

1. **Amplification of Pro-China and Irrelevant Content**: TikTok amplifies frontier influencers (travel and lifestyle content accounts) and irrelevant or clickbait material, to crowd out discussion of CCP-driven ethnic genocide and human rights abuses on its platform.
2. **Suppression of Anti-China Content**: TikTok's moderation algorithms significantly augment this suppression. **The views-to-likes ratio for anti-China content on TikTok was 87% lower than pro-China content even though the content was liked nearly twice as much.**
3. **Cross-Platform Influence Operations**: The CCP also uses frontier influencers and state-affiliated media to disseminate pro-China narratives to crowd out discussion of human rights abuses on Instagram and YouTube with tourism and culture content.
4. **Psychological Indoctrination**: A psychological survey of Americans (n=1214) shows that, among the platforms studied, TikTok screentime positively and uniquely predicted favorability towards China's human rights record. Notably, heavy users of TikTok (i.e., those with >3 hours of daily screentime) demonstrated a roughly 50% increase in pro-China attitudes compared to non-users. **This suggests that TikTok's content may contribute to psychological manipulation of users, aligning with the CCP's strategic objective of shaping favorable perceptions among young audiences**.
5. **Strategic Assessment**: NCRI assesses that the CCP is deploying algorithmic manipulation in combination with prolific information operations to impact user beliefs and behaviors on a massive scale and that these efforts prove highly successful on TikTok in particular. These findings underscore the urgent need for transparent regulation of social media algorithms, or even the creation of a public trust funded by the platforms themselves to safeguard democratic values and free will.

BACKGROUND

This study continues the NCRI's analysis of algorithmic bias. Therein, NCRI compared hashtag ratios between TikTok and Instagram for terms that are sensitive issues domestically and externally for the CCP. Though the study was preliminary, it concluded that there exists "a strong possibility that TikTok systematically promotes or demotes content on the basis of whether it is aligned with or opposed to the interests of the Chinese Government."[2] In the wake

[2] Network Contagion Research Institute. (2023, Dec 21). A Tik-Tok-ing Timebomb: How TikTok's Global Platform Anomalies Align with the Chinese Communist Party's Geostrategic Objectives. https://networkcontagion.us/reports/12-21-23-a-tik-tok-in-timebomb-how-tiktoks-global-platform-anomalies-align-with-the-chinese-communist-partys-geostrategic-objectives/

of publication, TikTok disabled its hashtag measurement functionality in a move that renders our December research findings irreplicable.

Given these platform changes undertaken in recent months, the methodological basis of this current study are *user journeys*. The user journey methodology recreates the on-platform experience of a newly created, organic user. As is discussed at length the following section, the accounts created for this study on TikTok, Instagram, and YouTube mimicked 16-year-old users located in the United States. These accounts were used to test the three platforms' search algorithm results when inputting keywords of hot-button issues for the CCP domestically. An advantage of user journeys over web scraping or paginated mass collection techniques is that manual operation of dedicated sock puppet accounts more closely reproduces the search results and content flow that an organic user is likely to encounter.

In recent years, user journeys are becoming an increasingly widespread technique for internet and social media algorithmic researchers. AI Forensics, a European non-profit research institute, partnered with Amnesty International in 2023 to conduct a user journey-driven analysis to study how the TikTok algorithm negatively influences teen mental health.[3] AI forensics' primary research design relied on automatic collection of post data via 40 sock puppets, paired with a manual audit that strove to heuristically confirm findings. The Amnesty International report concluded that through TikTok's algorithm, "children and young people are being exposed to a system which turns their psychological vulnerabilities into a means to maximize "user engagement."[4]

Following the October 7th attack by Hamas and the eruption of conflict across the Middle East, researchers with the Wall Street Journal (WSJ) embraced user journeys as a way to gauge algorithmic bias vis-a-vis the war in Gaza. The WSJ team created 8 sock puppet accounts posing as 13-year-old American teens and categorized conflict-related content that was served up as either pro-Palestinian or pro-Israel. The WSJ found that:

> "Similarly to other social-media platforms, much of the war content TikTok served the accounts was pro-Palestinian—accounting for 59% of the more than 4,800 videos served to the bots that the Journal reviewed and deemed relevant to the conflict or war. Some 15% of those shown were pro-Israel."[5]

This finding that four times as much pro-Palestinian conflict-centric content was served relative to pro-Israel content likewise coincides with the CCP's geopolitical interests in the Middle East, which have notably chilled towards Israel while warming towards the Palestinians and other

[3] AIForensics. (2023). Unmasking TikTok's influence on youth mental health: A comprehensive investigation. https://aiforensics.org/work/tiktok-kids

[4] Amnesty International. (2021). Driven into the darkness: How TikTok's 'For You' feed encourages self-harm and suicidal ideation (p. 64). https://aiforensics.org/uploads/POL_4073502023_ENGLISH_6ace7064be.pdf

[5] Schechner, S., Barry, R., Wells, G., French, J., Whitton, B., & Dapena, K. (2023, Dec 22). How TikTok brings war home to your child. The Wall Street Journal. https://www.wsj.com/tech/tiktok-israel-gaza-hamas-war-a5dfa0ee

international backers. Though TikTok has denied claims of anti-Israel bias, mounting research on its coverage of the ongoing Middle East conflict has sparked the interest of U.S. lawmakers.[6] The analysis undertaken in this current study assumes a similar experimental design to that of the WSJ though it focuses on domestic Chinese bellwether issues rather than global geopolitical concerns.

While user journeys represent a powerful tool for exposing algorithmic bias on TikTok, the scope of CCP digital influence operations is significantly broader and cross-platform in nature. One such influence vector the findings of this report illustrate is "frontier influencers," which are defined by the Australian Strategic Policy Institute (ASPI) as:

> "[An] increasingly popular set of YouTube accounts that feature mainly female China-based ethnic-minority influencers from the troubled frontier regions of Xinjiang, Tibet and Inner Mongolia, hereafter referred to as 'frontier influencers' or 'frontier accounts.'"[7]

As ASPI details, the CCP weaponizes *prima facia* benign travel vloggers to manage perceptions of China abroad, particularly as relates to restive border areas like Tibet, Xinjiang, and Inner Mongolia. As the report remarks, one tactic employed by these frontier influencers is ensuring "that state-backed content ranks well in search results because search-engine algorithms tend to prioritize fresh content and channels that post regularly."[8] As will be discussed below, our user journey findings directly confirm ASPI's 2021 findings.

Alongside the positive amplification performed by third-party frontier influencers and high-reach official CCP media accounts, hashtag hijacking is a technique that Beijing has employed to suppress undesired narratives. Hashtag hijacking is the deliberate attachment of trending hashtags or keywords to unrelated content or posts to 1) boost engagement for the network of bots performing the hijacking, and 2) suppress undesirable content bearing a target hashtag by flooding the social discourse with white noise (also bearing the target hashtag or keyword). A dramatic example of hashtag hijacking was on display during the CCP's 2022 crackdown on the protests against its zero-COVID policy that erupted across the mainland. As reported by The Guardian:

> "...Networks of coordinated bot accounts were targeting non-Chinese social media platforms to crowd out genuine posts about the demonstrations with spam content and by hijacking hashtags of names of Chinese cities. It said China's government was most likely to be behind the tactic."[9]

---

[6] Sung, M. (2023, Nov 14). TikTok disputes claims of anti-Israel bias amid calls to ban the app. TechCrunch. https://techcrunch.com/2023/11/14/tiktok-disputes-claims-of-anti-israel-bias-amid-calls-to-ban-the-app-palestine/

[7] Ryan, F., Impiombato, D., & Pai, H.-T. (2022, Oct 20). Frontier influencers: The new face of China's propaganda. Australian Strategic Policy Institute. https://www.aspi.org.au/report/frontier-influencers

[8] Ibid.

[9] Milmo, D. (2022, Dec 4). China accused of flooding social media with spam to crowd out protest news. The Guardian. https://www.theguardian.com/world/2022/dec/04/china-accused-of-flooding-social-media-spam-covid-protests

The intermingled operationalization of algorithmic manipulation on TikTok (top-down approach) alongside networks of frontier influencers, tourist vloggers, high-reach CCP-sponsored media assets, and hashtag hijacking on US-based platforms (bottom-up approach) should be viewed within the broader framework of the CCP's doctrine of influence and information warfare. As characterized by the French Armed Forces' Institute for Strategic Research, the CCP's information operations strive to achieve two primary objectives: 1) "seduce and subjugate foreign audiences by painting China in a positive light",[10] and 2) "infiltrate and constrain – a 'harsher' category of operations that do not involve seducing its opponents but rather bending them."[11] This study works to advance an understanding about the specific mechanisms driving the "seduce and subjugate" pillar of China's information warfare doctrine as observed across TikTok, Instagram, and YouTube around hot-button issues for Beijing.

METHODOLOGY

Collection Methodology

The objective for user journey data collection was to collect ±300 videos for each of four target keywords (i.e., "Uyghur", "Xinjiang" "Tibet", "Tiananmen") across three different social media platforms (i.e., TikTok, YouTube, Instagram), producing ±3435 results total, as some search feeds stopped serving content before 300 videos per term was reached. These keywords were selected given their importance in the CCP's information warfare and propaganda doctrine, which enshrines projecting a positive image of China both inwards and outwards as a core objective of its national security agenda. Results on TikTok, a subsidiary of Chinese tech giant ByteDance, was compared to those on Instagram and YouTube, owned by U.S. companies Meta and Google, respectively, in order to measure how content served up by the search algorithm differs between platforms.

The user journey methodology strives to recreate the organic experience a generic new account would undergo. All accounts were created with IP addresses located in the USA. Equal numbers of male versus female accounts were created, and all accounts were marked as 16 years of age to best mimic the organic user experience of an American teenager. Both TikTok and Instagram collection was performed on mobile Android phones and recorded using a phone screen recording app called V Recorder, while YouTube collection was performed on a desktop computer and recorded via a screen recording tool.

In total, 24 accounts were created for the purposes of this study. A new account was created for each keyword for each platform to ensure experimental sterility by restricting the personal identifiers visible to the platform algorithms to 1) IP location and 2) user age. Beyond account creation, searching for the target keyword, scrolling through video results, and saving/bookmarking viewed content, no additional actions were performed that could skew the

[10] The Institute for Strategic Research. (p. 413). https://www.irsem.fr/en/index.html
[11] Ibid.

profile's search preferences (e.g., no accounts were followed, no prior searches were performed, no engagements except views and saves were performed).

A standard collection methodology was followed for all keywords across each platform. In conducting the TikTok data collection, the user began by typing the term into the Search field and clicking on the first video that appeared. Subsequently, the user scrolled through each subsequent video, saving each one. Every video was played for at least 15 seconds or until the video concluded. Upon completing the recording session, the user navigated to the Saved page on the User Profile to locate all the saved videos from the session. The upload date and the link for each video were then copied into a spreadsheet.

For Instagram data collection, the user started by entering the term in the Search field and selecting the first post that appeared. The user then scrolled through each subsequent post, saving each one. Videos were played for at least 15 seconds or until they finished, and posts containing multiple images were fully viewed by swiping through each image. After recording, the user accessed the Saved page on the User Profile to retrieve all saved posts from the session. The upload date and the link for each post were copied into a spreadsheet.

During YouTube data collection, the user entered the term into the Search field and hovered over each video in the list, allowing it to play for 15 seconds (excluding YouTube Shorts and videos that are in playlists). After completing the recording, the user scrolled back to the top of the list and clicked on each video to copy the upload date and URL into a spreadsheet.

Coding Methodology

Following data collection, the first phase of analysis categorized content as either pro-China, anti-China, neutral, or irrelevant. The search terms analyzed are inherently political for the CCP, facilitating a clearer and more intuitive coding process than would be possible with broader, more neutral terms such as "China" or "democracy".

Blind coding[12] was performed by two human analysts. In cases where there was a disagreement in the coding determination, a third subject matter expert (SME) was tasked with arbitrating and assigning a final coding category. Given the subtlety of messaging alongside the intermixture of textual, visual, and audio semiotics that were assessed in making a coding determination, we believe that human analysis tends to be more accurate and granular than AI or machine-driven coding.

Despite reliance on human analysts for content coding, our average margin of intercoder disagreement across all platforms and keywords was remarkably small. The average intercoder disagreement across all keywords and platforms was measured as 12.81%, with a range of

[12] In this context, blind coding connotes that the analysts categorizing video results were isolated from one another to prevent conformity bias between their designations. Furthermore, this research was hypothesis agnostic in order to minimize coder bias.

18.9% (3.8% to 22.7%). Specific values for each keyword and platform are displayed in the following table:

Intercoder Disagreement	Tibet	Tiananmen	Uyghur	Xinjiang
TikTok	13.5%	3.8%	7.7%	9%
Instagram	16.3%	14%	8.3%	20.3%
YouTube	21%	12%	4.1%	22.7%

Table 1: Intercoder Disagreement in Content Coding

Our analysis utilized a systematic approach to coding video content, ensuring consistency and accuracy in categorizing videos. The coding parameters were customized for each keyword, and served as a blueprint for human analysts responsible for the process:

Xinjiang

- Pro-China: Official promotional content, frontier influencer content, showcasing of minorities' folk customs while whitewashing cultural erasure, idyllic portrayals of rural life, claims of Western narrative misrepresentation
- Anti-China: Content highlighting Uyghurs' plight in China, calls for boycotts of Chinese products grown in Xinjiang, Chinese human rights abuses and suppression of internal dissent
- Neutral: Personal photos, informational graphics, unbiased news reports, historical artifacts, consumer goods
- Irrelevant: Content unrelated to Xinjiang, apolitical Xinjiang diaspora content

Uyghur

- Pro-China: Highlight Uyghur/Xinjiang folk culture (food, dance, dress, women), frontier influencers exploring Xinjiang/Uyghur heartland
- Anti-China: Content highlighting Uyghurs' plight in China, unlawful detention, cultural erasure, suppression of civil liberties, etc.
- Neutral: Diasporic communities, apolitical Uyghur-language songs or media, professional travel photographers and/or Western tourists unaffiliated with official operators.
- Irrelevant: Content unrelated to Uyghurs

Tibet

- Pro-China: Official promotional content, state-registered tourism companies, frontier influencer content, idyllic portrayals of rural life, echoing the CCP narrative points that Tibet has been liberated
- Anti-China: Mentions of Tibetan liberation, coverage of the exilic government and political statements from the Dalai Lama, videos containing politically charged hashtags like #freetibet, #SaveTibet and #proudtibetan, footage of protests, and details of Tibetan cultural erasure by the CCP
- Neutral: Informational presentations, unbiased historical content, coverage of Tibetan Buddhism, its rituals and material culture
- Irrelevant: Content unrelated to Tibet, reactions to Tibetan culture, and Tibetan consumer & folk-art products

Tiananmen

- Pro-China: Patriotic songs, official travel promotions, flag raising and other nationalist events, denials of the massacre and revisionist historical takes, scenic pictures of the square that bear no mention of the massacre
- Anti-China: Condemnations of the massacre, commemorations worldwide of the massacre by victims and dissidents, "Tank Man" imagery, memes highlighting the event
- Neutral: News coverage of worldwide anniversaries for the massacre, tangential mentions of Tiananmen
- Irrelevant:[13] Content unrelated to Tiananmen Square or the 1989 massacre

DATA ANALYSIS

User Journey Data

User journey data shows patterns consistent with the conclusion that TikTok is suppressing anti-China content in order to diminish the reach of narratives critical of the CCP (see Figure 1). For all four search terms examined, TikTok contained the lowest proportion of anti-China content compared to Instagram of YouTube. Simultaneously, we observed that pro-China and irrelevant content constituted between 61%-93% of search results on TikTok, illustrating how flattering or distracting content is systematically amplified as a way to subsume the low proportion anti-China coverage extant on the platform.

[13] Note that 8 videos (<3%) served for “Tiananmen” on TikTok contained scenic content of the Tianmen mountains; these were categorized as irrelevant.

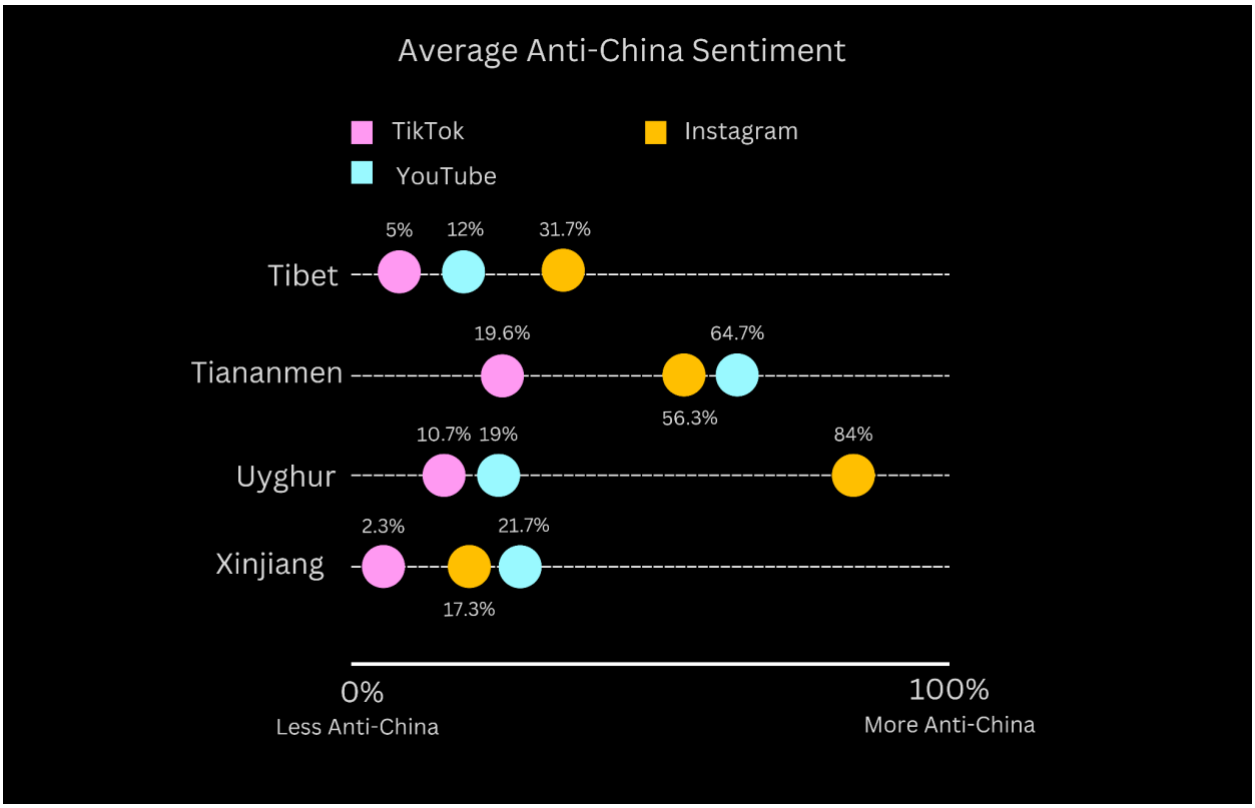


Figure 1: Summary of Anti-China Bias

The results of the user journeys performed for “Tiananmen” (Figure 2) support a hypothesis that TikTok’s search algorithm is suppressing anti-China content while amplifying pro-China and irrelevant content, relative to Instagram and YouTube. For instance, 26.6% of search results on TikTok for “Tiananmen” were categorized as pro-China, versus 16.3% on Instagram and a mere 7.7% on YouTube.

Conversely, the proportion of anti-China content on YouTube (64.6%) and Instagram (56.3%) was more than double that observed on TikTok (19.6%). It is also significant that 45.6% of content served on TikTok was flagged as irrelevant compared to <10% for both Instagram and YouTube. This reflects the “seduce and subjugate” strategy outlined in the introduction where content showing other Chinese cities and natural attractions was used to subsume mentions of Tiananmen.

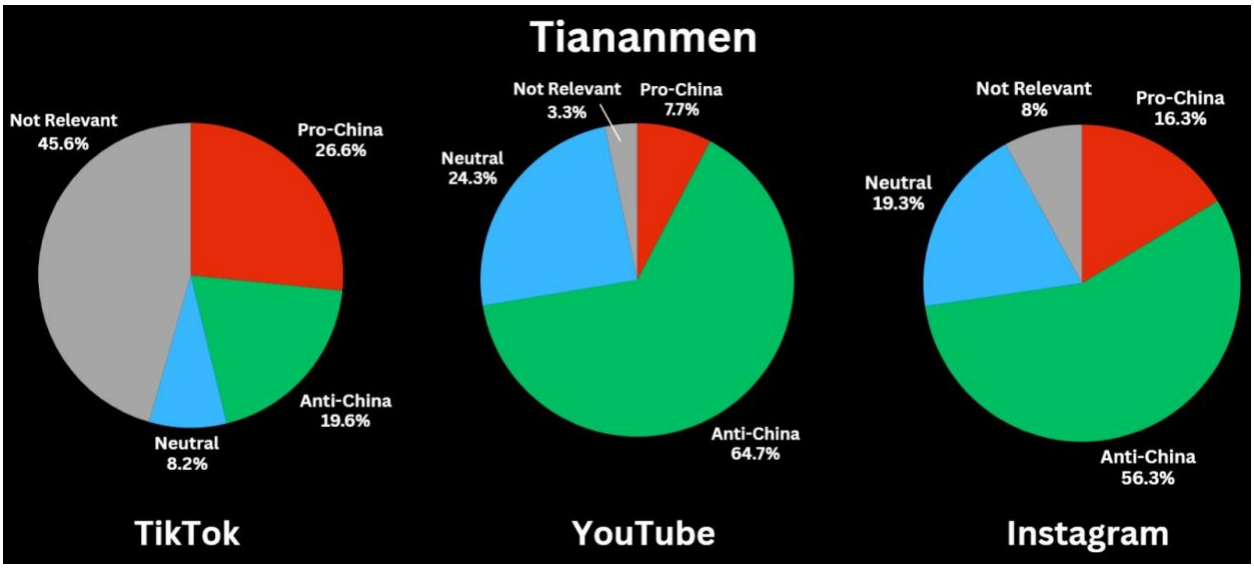


Figure 2: "Tiananmen" User Journey Summary

Broadly speaking, a similar breakdown of coding categories was observed for the search term "Tibet" (Figure 3). TikTok contained the highest proportion of pro-China content (albeit only slightly more than Instagram) and the lowest proportion of anti-China content by far across all three platforms. Analysis also showed that the proportion of neutral content was similar across all three platforms, which we attribute to apolitical mentions of Tibet in relation to Tibetan Buddhism and spirituality, a subject that enjoys widespread familiarity among the general public in the West. Overall, however, these findings, again testifies to TikTok's proclivity to suppress anti-China content and amplify pro-China content, relative to Instagram and YouTube.

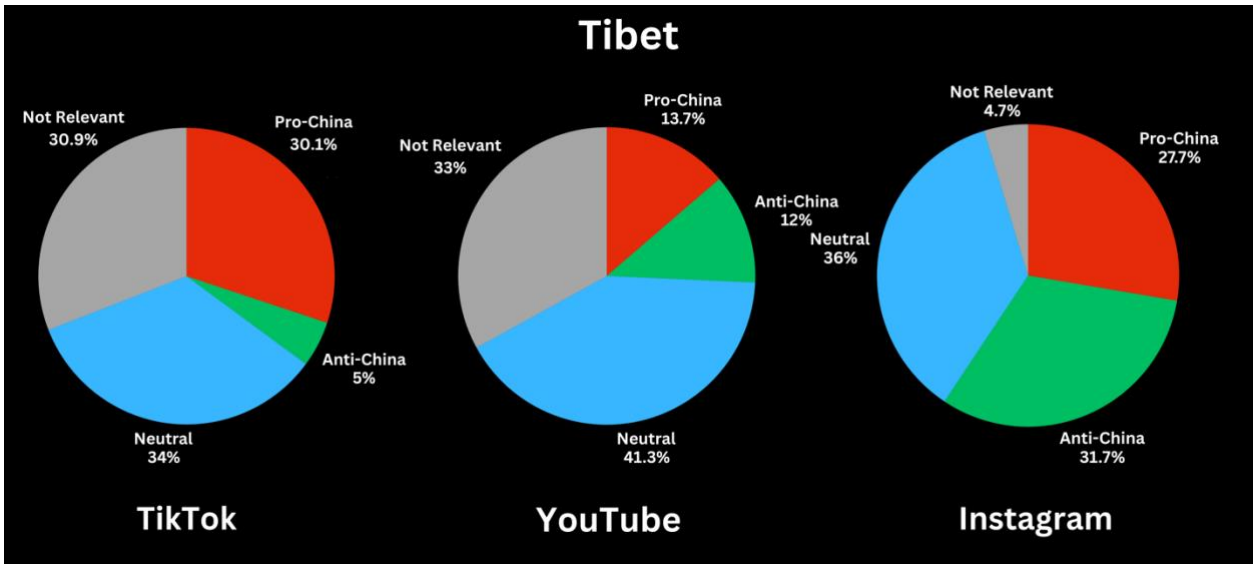


Figure 3: "Tibet" User Journey Summary

For the search term “Uyghur” (Figure 4), anti-China content was measured at a mere 10.7% of TikTok results vs. 19% of YouTube and an overwhelmingly 84% of Instagram results, thereby testifying to the hypothesis that TikTok is suppressing anti-China content. In contrast, pro-China content comprised 17% of TikTok data compared to 2.7% of Instagram data yet an unintuitive 49.2% of YouTube results. While at first glance this increase in pro-China content on YouTube relative to TikTok appears unintuitive, upon closer examination these differences can be attributed to the strategic cultivation of pro-China media assets on YouTube and Instagram. We observed that these assets have effectively dominated on-platform search results in what constitutes an example par excellence of bottom-up influencing on YouTube and Instagram.

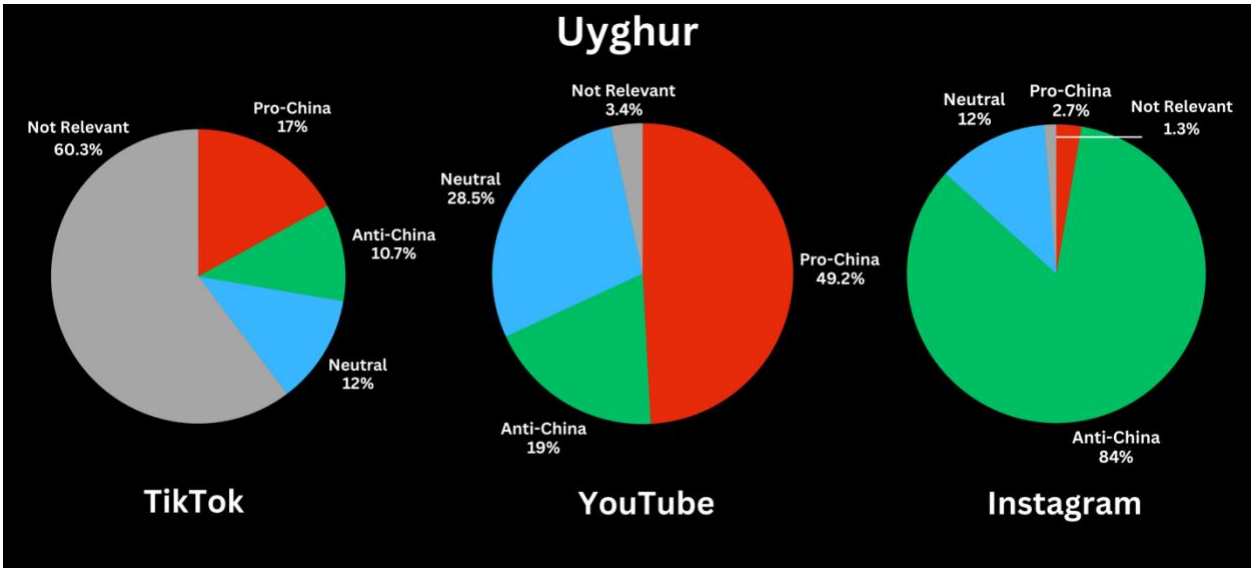


Figure 4: “Uyghur” User Journey Summary

This anomalously high proportion of pro-China content on YouTube was driven by the fact that 40% of the total content collected emanated from a single account, @uyghurbeauty. Manual coding of @uyghurbeauty resulted in categorization as pro-China given its modus of amplifying folk culture as a means of whitewashing the ongoing Uyghur cultural genocide in Xinjiang. An in-depth analysis of @uyghurbeauty and similar accounts that are playing a core role in hashtag keyword/hijacking of discourse around the search term “Uyghur” is presented below in Appendix I.

Figure 5: @Uyghurbeauty YouTube Profile

Also of note in the “Uyghur” user journey data set was that 60.3% of all content served by TikTok’s search algorithm was flagged as irrelevant in contrast to <5% for Instagram and YouTube (see Figure 6). This differential demonstrates how the CCP influencing tactic of hashtag/keyword hijacking being effectively operationalized in service of the “seduce and subjugate” strategy.

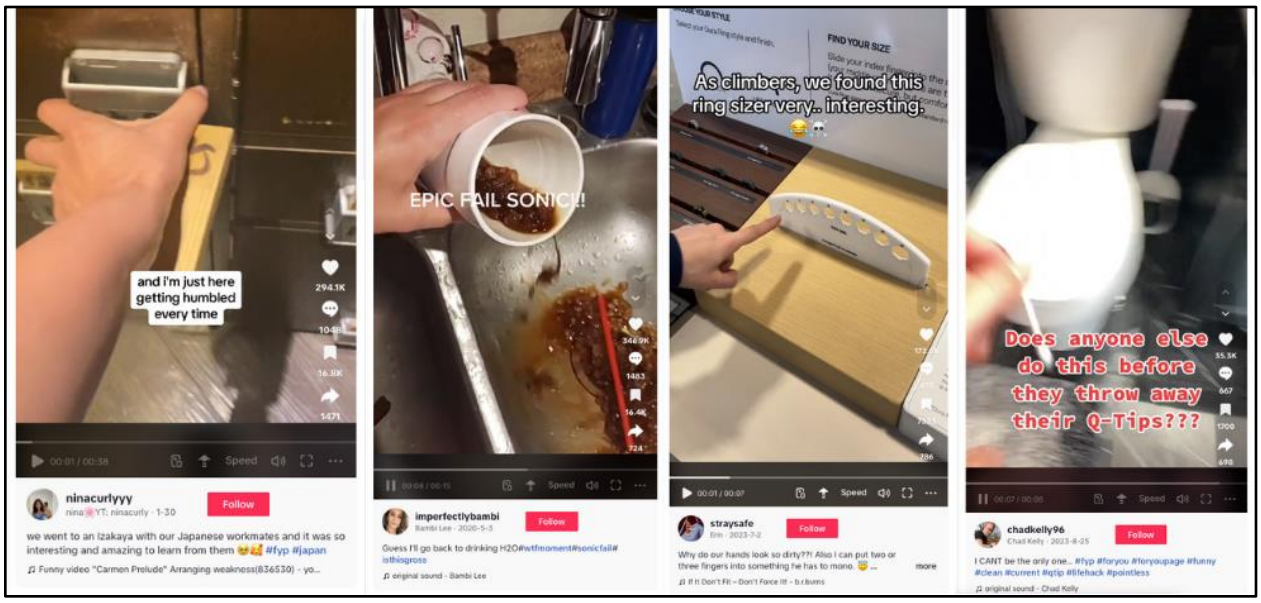


Figure 6: Irrelevant Content Served for “Uyghur” on TikTok

The user journey data profile for “Xinjiang” (Figure 7) resembles that of “Uyghur” in some key respects. On TikTok, we found a virtual absence of anti-China content (2.3%) compared to 21.7% on YouTube and 17.3% on Instagram. Also of note is the high proportion of irrelevant content served to users on TikTok for “Xinjiang,” which was measured at 69.4% of all results. This finding supports the hypothesis that the TikTok search algorithm systematically

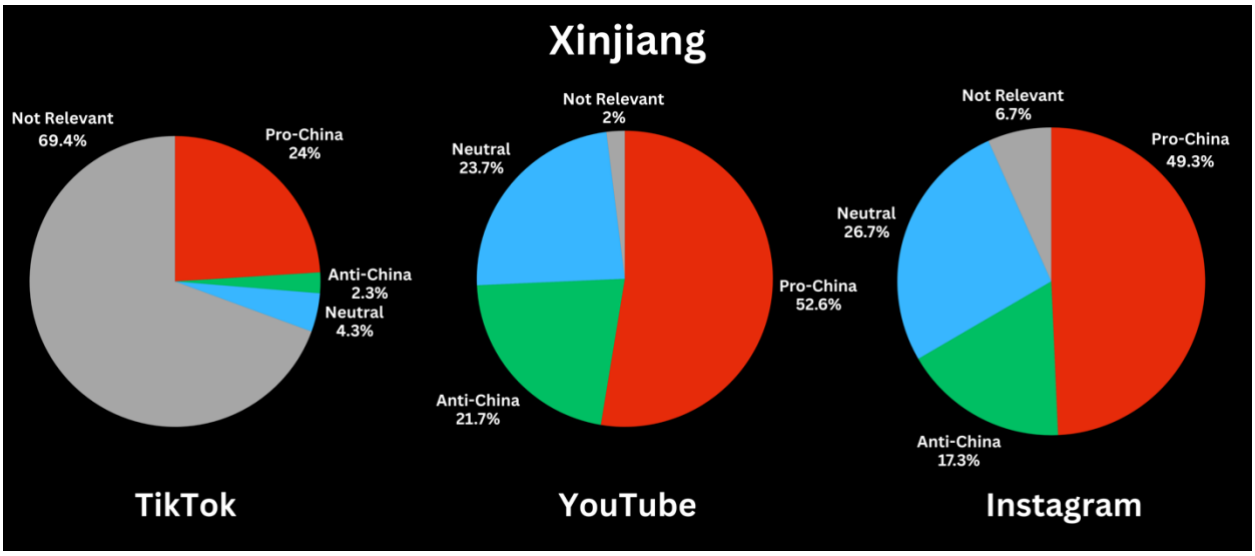


Figure 7: “Xinjiang” User Journey Summary

suppresses undesirable anti-China content while flooding search results with irrelevant content, much of which appeared to be generic clickbait or tourist content portraying other cities or regions of China in a positive or attractive light.

In contrast to the user journey data on the other keywords examined, for “Xinjiang,” a user was served more than twice as much pro-China content on both YouTube (52.6%) and Instagram (49.3%) than on TikTok (24%). As was the case of “Uyghur” on YouTube, we attribute this anomalous finding on Instagram and YouTube to a combination of pro-CCP media assets—both frontier influencers and official media accounts-- that have successfully dominated the social discourse around these keywords using a bottom-up approach to subsume search results.

On YouTube, four of the top five most active accounts that figured in search results were official or semi-official CCP media sources (CGTN, South China Morning Post, ShanghaiEye, and CCTV Video News Agency; see Figure 8) that enjoy an accumulated subscriber pool of 8M users. These four users generated 21.7% of all content retrieved for “Xinjiang” and nearly 40% of all pro-China content on YouTube, thereby testifying to the extent they have successfully dominated search results for “Xinjiang.”

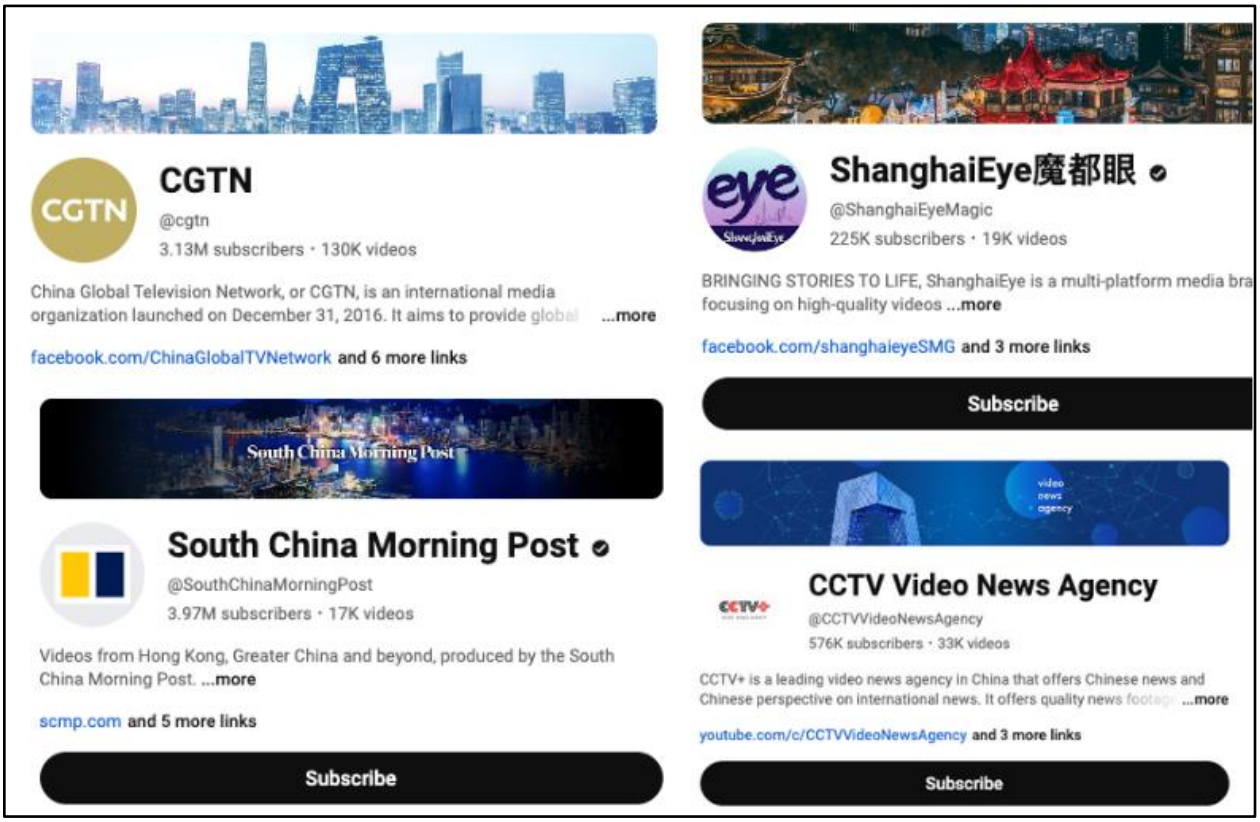


Figure 8: Pro-China Media Assets on YouTube

In a similar vein, on Instagram the most active account is @wilchinawildchina, which is based in China and appears to be a frontier influencer with nearly 68K followers. An unofficial fan page (@dilireba_dilmurat) for the Uyghur-Chinese model and CCP brand ambassador Dilraba Dilmurat

ranked as the fourth most active poster, and official state media @cgtn as the fifth most active poster. The high visibility enjoyed on Instagram and YouTube by either official CCP media sources or affiliated third parties around discussion of “Xinjiang” testifies to the successful cultivation and weaponization of social media assets to again achieve the “seduce and subjugate” pillar of Beijing’s influence doctrine.

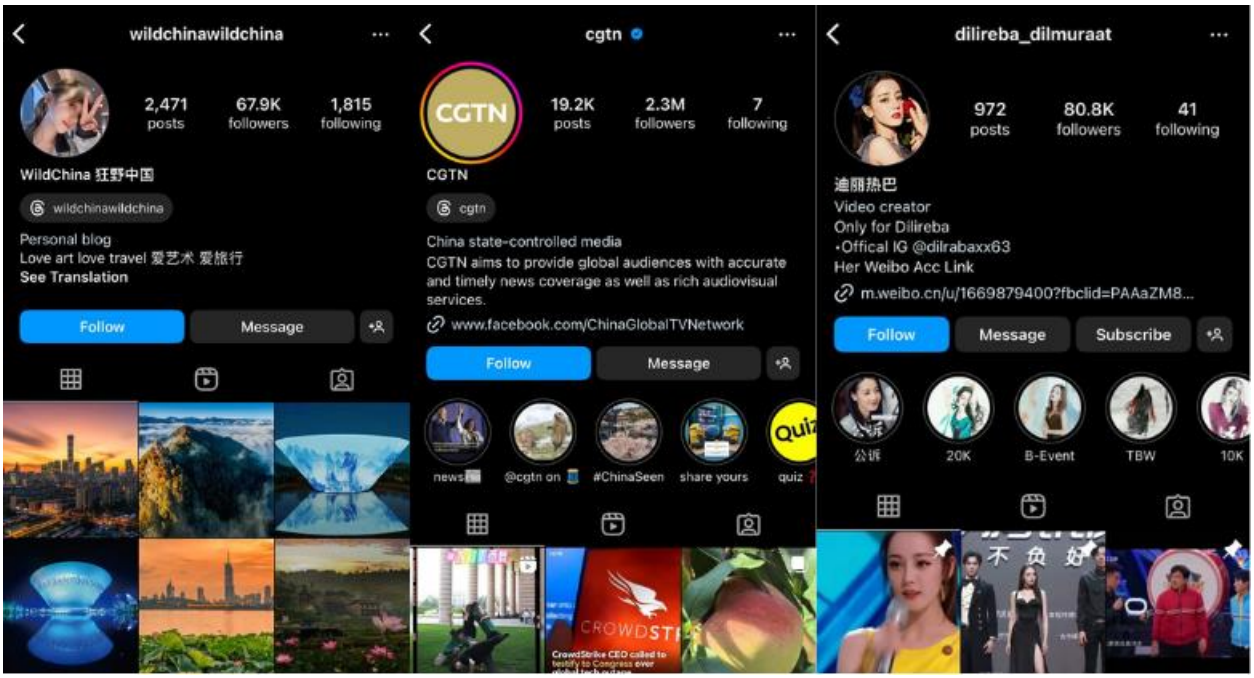


Figure 9: Pro-China Media Assets on Instagram

Engagement Metadata

Whereas the first prong of our data analysis shed light on the pro-China bias of TikTok’s search algorithm, engagement metadata analysis[14] illuminates which category of content (pro-China vs. anti-China) accrued more engagement (likes and views) from platform users. This allows us to determine not only what category of content was served most frequently, but what category of content was most popular with other users.

As has been widely documented by researchers and marketing experts, the initial engagement awarded to videos posted on TikTok (views, full-length plays, likes, comments, shares) influences the extent to which the algorithm exposes the content to successively broader audiences.[15] Despite this accepted rule of thumb, our findings point to the opposite conclusion than is suggested by the user journey data presented above, namely that though pro-China content is served up by the TikTok search algorithm much more often, anti-China content appears far more engaging for organic users.

[14] See Appendix III for full metadata engagement metrics. In this study, engagement metadata refers to the engagements (views, likes, comments, shares, etc.) awarded to videos analyzed.

[15] Klug, D., Qin, Y., Evans, M., Zhang, B., & Zhang, K. (2022). An empirical investigation of personalization factors on TikTok. *arXiv*. https://ar5iv.labs.arxiv.org/html/2201.12271

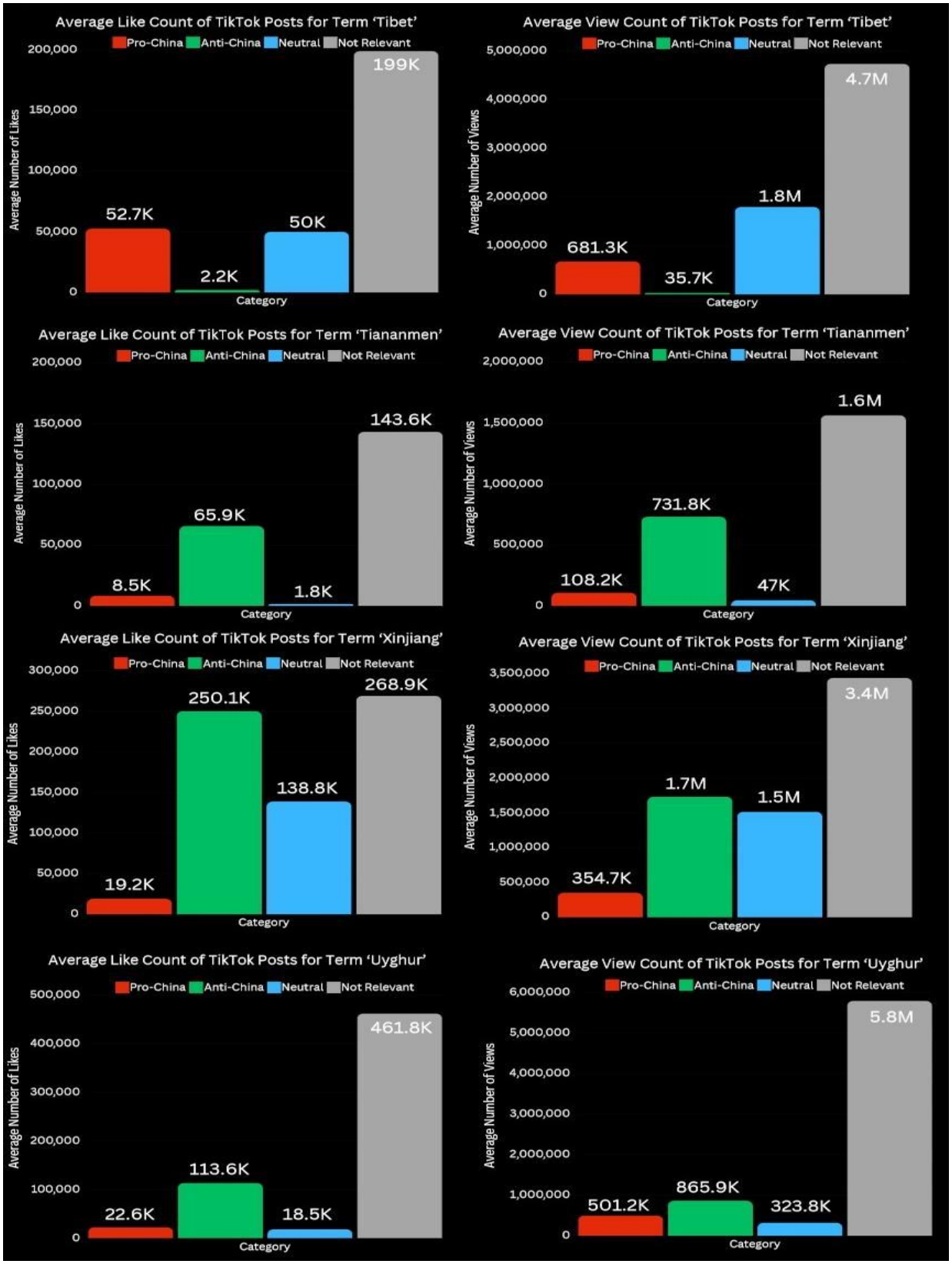


Figure 10: TikTok Engagement Metadata

For three of the four search terms collected on TikTok (excepting Tibet), engagement awarded to anti-China content was consistently and significantly higher than that given to pro-China content. This finding suggests that notwithstanding the indications of systematic suppression of anti-China content on TikTok elucidated above, said anti-China content accrued significantly higher average likes and views than pro-China content (see Figure 10).

Engagement metadata for Tibet on TikTok showed an inverse finding relative to “Tiananmen,” “Uyghur,” and “Xinjiang” in that pro-China content accrued significantly higher average views and likes than anti-China content. We partially attribute this skew to a single, highly viral pro-China TikTok video collected for “Tibet” that received 19.4M views and 2M likes. As indicated in the table, this video’s engagement metrics are orders of magnitude above the average observed across all four search terms and skewed TikTok Tibet results in a markedly pro-China direction.

Interestingly, the account (@rawangm1; see Figure 11) that uploaded the viral clip fits the profile of a classic frontier influencer—"mainly female China-based ethnic-minority influencers from the troubled frontier regions...[who project] a false sense of legitimacy and transparency about China’s frontier regions that party-state media struggle to achieve.”[16] The account also shares a link to an ecommerce landing page for Chinese-themed consumer products. Given the lack of connection between the product portfolio displayed on the landing page and Tibet or Tibetan culture, it is possible that the account is leveraging the algorithmic bias that amplifies viral frontier influencer content to drive affiliate marketing traffic.

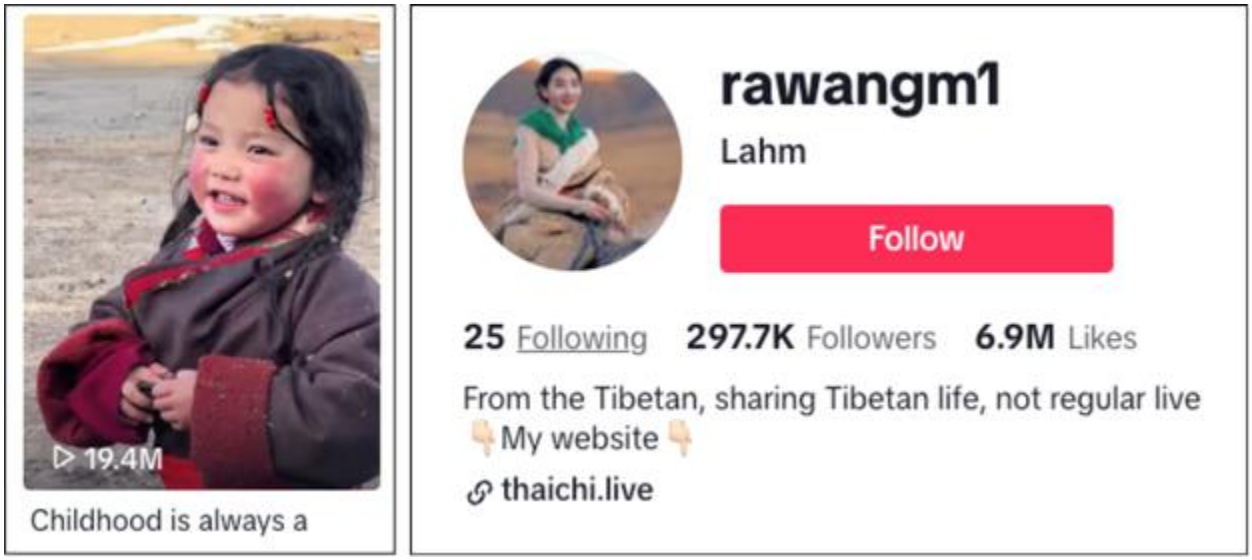


Figure 11: Example of Tibetan Frontier Influencer

Returning to the engagement metadata analysis, the highest average likes and views across all four search terms analyzed on TikTok was granted to irrelevant content, signifying that generic clickbait content consistently attracts higher engagement than politicized content from users

[16] Ryan, F., Impiombato, D., & Pai, H.-T. (2022, Oct 20). Frontier influencers: The new face of China's propaganda. Australian Strategic Policy Institute. https://www.aspi.org.au/report/frontier-influencers

(see Figure 12). **In sum, TikTok engagement metadata analysis shows that despite TikTok search algorithm's tendency to amplify pro-China content and suppress anti-China content:**

- **Anti-China content engenders significantly higher user engagement than pro-China content.**
- **Apolitical (i.e., coded as irrelevant) clickbait receives the overall highest engagement for all categories measured.**

Analysis of engagement metadata for YouTube and Instagram yielded higher average likes and views (YouTube only) per video for anti-China content. As was observed on TikTok, "Tibet" remained an outlier on Instagram, and to a lesser extent, on YouTube. The results observed on YouTube show that a higher average number of likes and views were awarded to anti-China content across all search terms, the sole exception to this finding being 60K less views for anti-China content mentioning Tibet relative to pro-China content, as shown in the chart below. Also of note is that for "Tibet," "Xinjiang," and "Uyghur," irrelevant content accrued a higher average of likes and views than either pro- or anti-China content. These findings parallel those observed on TikTok, and suggests that on video-centric platforms (YouTube & TikTok), generic clickbait content draws higher user engagement than politicized content.

On Instagram, average likes for pro-China content mentioning Tibet were 1.8x higher than average likes awarded to anti-China content. That said, for the other three search terms analyzed (Tiananmen, Xinjiang, Uyghur), anti-China content received 2.8x, 40.8x, and 8.3x more average likes, respectively, than pro-China content. Though Instagram engagement metadata examined herein was limited to likes (as views for post data is not accessible), the overall results closely mirrored that observed on TikTok, i.e., anti-China content generally garners significantly higher user engagement than pro-China content.

We found that though the ratio of average views per likes on TikTok was nearly twice as high for pro-China content (15.41 vs. 8.22), for three out of the four search terms, average views and likes were significantly higher for anti-China content. **The key takeaway from this ratio is that despite the TikTok search algorithm's pro-China bias, users favored anti-China content with likes and views over pro-China content in three out of four keywords analyzed (see Appendix III).**

In terms of average views per like across all search terms for YouTube, pro-China content received a value of 44.69 whereas anti-China content received 39.16. As was the case for TikTok, pro-China content received a higher ratio of views per like than anti-China content, notwithstanding the fact that anti-China content received significantly higher total engagement nearly across the board.

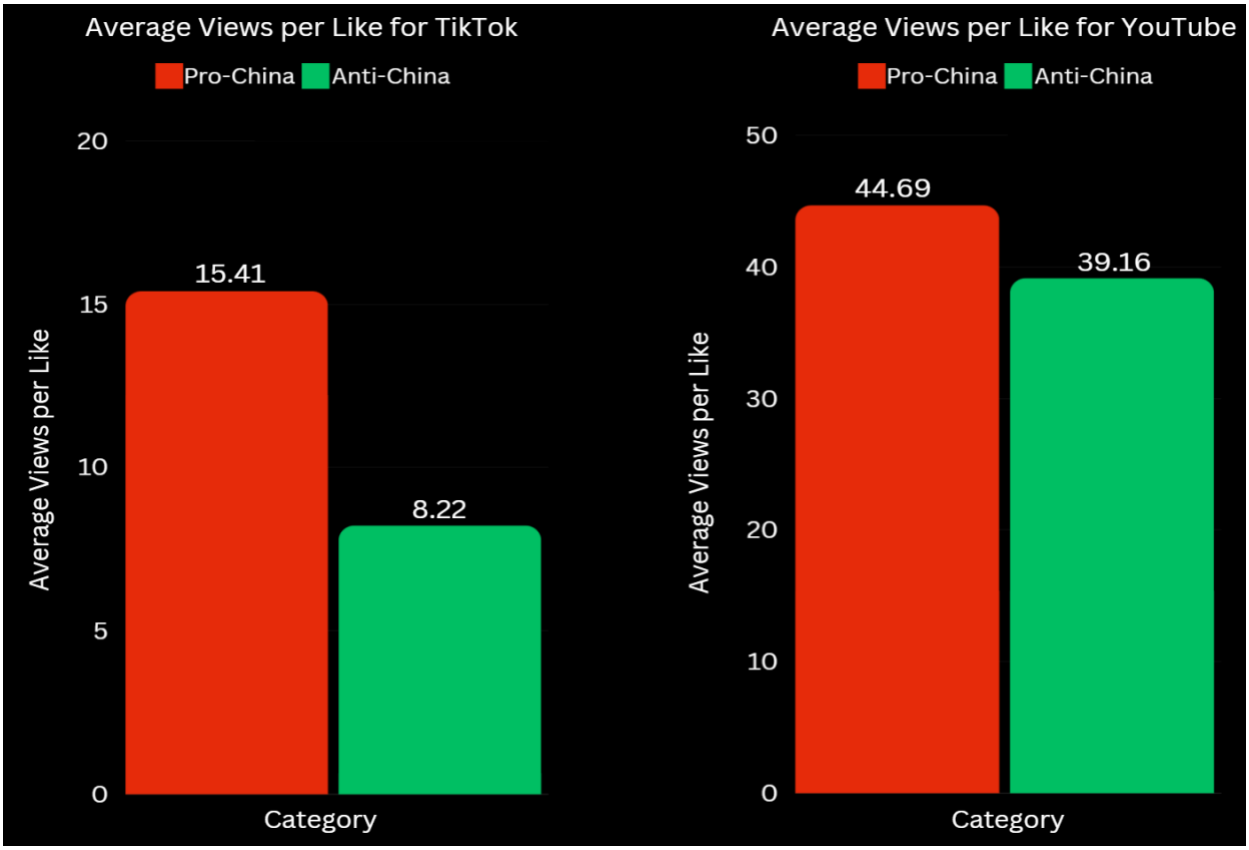


Figure 12: Average Views per Like for TikTok and YouTube

Whereas the views-per-likes ratio for pro-China content on TikTok (15.41) was 87% higher than anti-China content (8.22), the same ratio on YouTube was just 14%. This indicates that despite anti-China content receiving higher engagement (views, likes) than pro-China content on average, the algorithmic bias on TikTok drives more views per like to pro-China content than anti-China content.

This finding seemingly contradicts the previously cited finding (footnote 15) that more popular content (i.e., more likes) drives higher exposure and more views. **Therefore, it appears that TikTok's algorithm uniquely favors pro-China content in terms of views per like, irrespective of the overall engagement levels. Moreover, this finding is significantly more pronounced on TikTok than YouTube, suggesting a TikTok-specific bias.**

SURVEY DATA[17]

Initial results of the above research suggest that TikTok possesses a drastically smaller proportion of anti-CCP material relative to both YouTube and Instagram and also suggests that anti-CCP material is being manipulatively suppressed from view, either algorithmically or through moderation. But these findings do not prove that manipulated changes in platform content might be reflected in user attitudes and psychology. This is an important point to elucidate because if surreptitious manipulation of social media platforms undermines user autonomy, this creates significant threats to the capacity for free thought and speech.

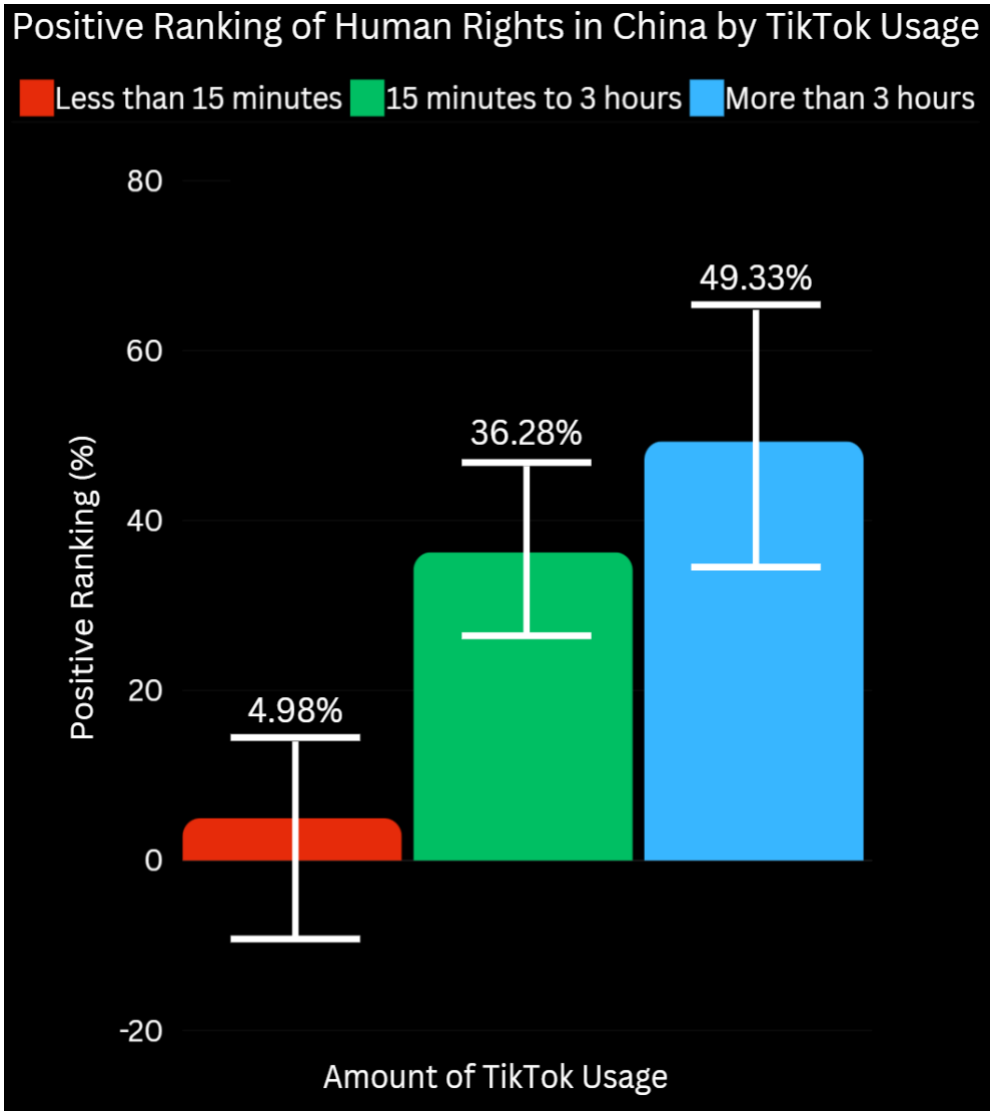


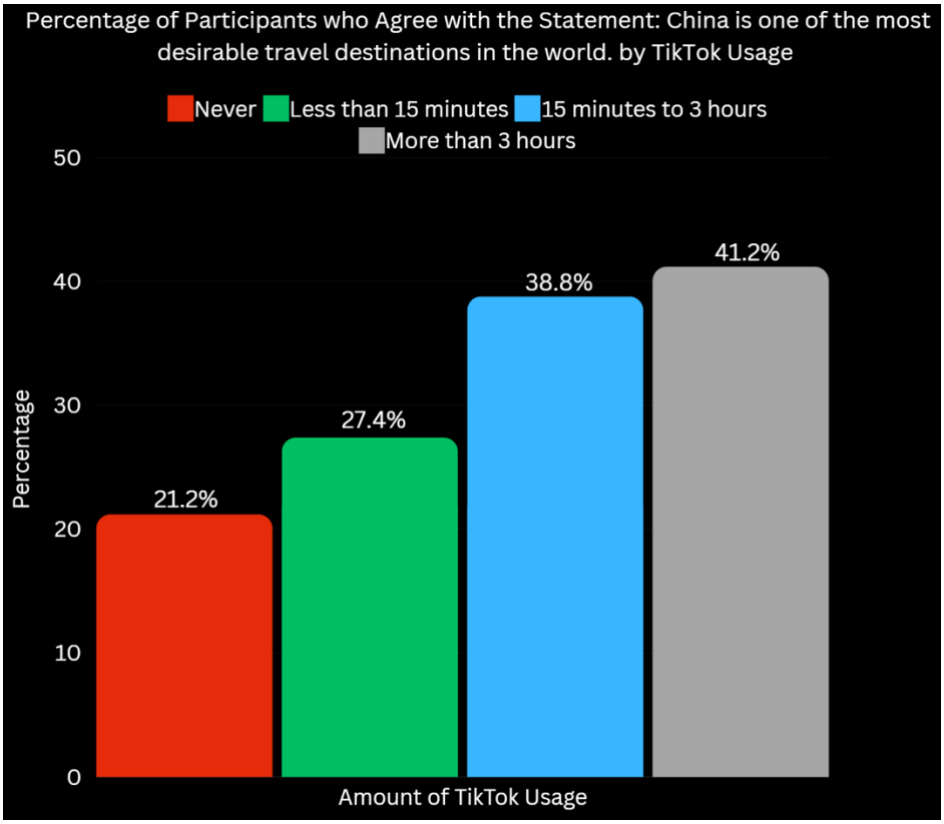
Figure 13: Human Rights Ranking in China by TikTok Usage

[17] The full survey data set, results, and methodology can be found as an accompanying addendum on the report page. Furthermore, full details on sampling methods, exclusion criteria, and analysis can all be accessed in said addendum.

To better understand the relationship between mass manipulation of content and potential impacts on underlying user psychology/attitudes, NCRI ran a survey through Amazon's Prime Panels Service (n=1214) wherein respondents reported on their daily social media platform use. Survey questions were developed to probe positive/negative perceptions of China's human rights records, as well as subjects' feelings about China as a travel destination (informed by NCRI's analysis of putative CCP information operations which strongly featured travel/culture influencers). To disguise the purpose of the survey from subjects, these questions were randomized with similar questions regarding numerous countries.

When controlling for age, among the video-streaming platforms we studied, TikTok use significantly and uniquely predicted more positive perceptions of China's human rights record ($\beta = .21$, $t(4.51)$, $p < .0001$; see Figure 13 and Appendix II). This was particularly notable for heavy users (those who use 3 or more hours on the platform daily) who showed a 49% increase in positivity towards China's human rights records relative to non-users. By contrast, use of YouTube and Instagram showed no significant relationship on users' perception of China's human rights record.

Similarly, TikTok use, unlike Instagram and YouTube use, significantly predicted more positive perceptions of China as a travel destination ($\beta = .17$, $p < .0001$; see Figure 14 and Appendix II) and further, predicted users would agree with the statement that "Tiananmen square is mostly known for being a popular travel destination" ($\beta = .18$, $p < .0001$; see Appendix II). These findings were also especially notable for heavy users, with such users showing a 41% increase



in favorability to China as a travel destination and 48% increase in agreement with the notion that Tiananmen square is mostly known as a tourist site.

Figure 14: China Favorable Travel Destination by TikTok Usage

CONCLUSIONS AND RECOMMENDATIONS

This report establishes that TikTok algorithms actively suppress content critical of the Chinese Communist Party (CCP) while simultaneously boosting pro-China propaganda and promoting distracting, irrelevant content. Through the use of travel influencers, frontier lifestyle accounts, and other CCP-linked content creators, the platform systematically shouts down sensitive discussions about issues like ethnic genocide and human rights abuses. These insidious tactics have far-reaching consequences that psychological survey data show are effective in shaping the views of its intended target: the minds of young users. These users, through targeting or information environments engineered to sublimate free speech, appear to absorb these biased narratives unwittingly, leading to a distorted understanding of critical global issues.

The tension between free speech and national security has become a prominent issue in the digital age, particularly with platforms like TikTok. First Amendment advocates argue that banning such platforms infringes on free speech, while national security experts highlight the national security risks posed by potential foreign influence. Our research indicates that these two concerns are not mutually exclusive. Free speech can be abridged, and democratic institutions destabilized, as effectively through algorithmic manipulation of carefully indoctrinated masses as by direct censorship. Our research highlights how algorithmic manipulation is undermining free expression and destabilizing democratic nations from within.

Taken together, these findings underscore the urgent need for developing robust methods to pressure test algorithms and detect when they subvert free expression without user consent. The recent work by the Institute for Strategic Dialogue (ISD), which exposed how TikTok's algorithms can inadvertently amplify harmful content, which is a step in this direction. Accordingly, we recommend the creation of a Civic Trust funded by both platforms and the public that will be empowered to develop a systematic pressure testing system to set standards and ensure transparency. This system would help identify when platforms are manipulating user perceptions, forcing individuals to adopt ideas they did not choose, often without their awareness. Such a framework is essential to protect the integrity of free speech and the sanctity of free will; in doing so, it will maintain the democratic values within which these platforms operate, at least in free societies.

Furthermore, this Civic Trust must also develop fair and common consequences for platforms found to be undermining free expression. If social media algorithms are found to be subverting the very democracies that provide them the freedom to operate, they are both unjust and dangerous. There must be accountability and corrective measures to ensure that platforms are not exploited by state actors to erode democratic institutions and values.

LIMITATIONS

While our study provides significant insights into the manipulation of content on TikTok and other social media platforms, there are inherent limitations that should be acknowledged. First, our research primarily relies on the analysis of content served to newly created, organic accounts. While this methodology is designed to mimic the experience of typical users, it does not account for personalized content that may be delivered based on individual user histories and interactions over time. Consequently, the data may not fully capture the breadth of content suppression or amplification experienced by all users.

Additionally, the coding and classification of content as pro-China, anti-China, or irrelevant involved subjective judgment, even with the use of blind coding and a third-party arbitrator for discrepancies. Although efforts were made to minimize bias, the potential for interpretative differences remains. Furthermore, our study did not explore the full range of user engagement metrics, such as comments and shares, which could provide deeper insights into user reactions and the spread of content across networks.

Lastly, a clear limitation in this report, which compared video streaming platforms, was that it was not a fully comprehensive cross-platform analysis. Our findings highlight a novel and concerning shift from merely observing evidence of content influence on TikTok to detecting psychological changes among its users because TikTok users not only encounter biased content but as hypothesized, also exhibit altered perceptions and attitudes in favor of the CCP. More comprehensive efforts are needed to better understand the relationship between user psychology and the prevalence of manipulated content.

APPENDIX I

OSINT Analysis of @UyghurBeauty

In the user journey performed for "Uyghur," we observed that 40% of all video content collected on YouTube was uploaded by a single channel, @uyghurbeauty. On YouTube, @uyghurbeauty has 96K subscribers, a total of nearly 50M aggregate views, and over 3K video uploads. The channel description is a generic "Discover the beauty of Uyghur culture and traditions," and a link is provided to an identically named Instagram profile, @uyghurbeauty, with 161K followers.

The @uyghurbeauty YouTube channel is based in the United States, while the Instagram's location is not specified. Most of the recent content is highly formulaic, characterized by a music video or folksong with the following cookie-cutter text appended in the description:

> "We share Uyghur folk songs and Uyghur traditional dance. You will learn about Uyghur people and enjoy the beauty of Uyghur culture. Uyghur Beauty Channel operates exclusively on YouTube. Please don't forget to hit the like button and subscribe to see more!"

On Instagram, a similar content pattern is observed alongside the addition of a largely static list of hashtags that include: #uyghur #uyghursong #uyghurbeauty #uyghurgarden #uighur #uyghurpeople #uyghurtraditions #uyghurculture.

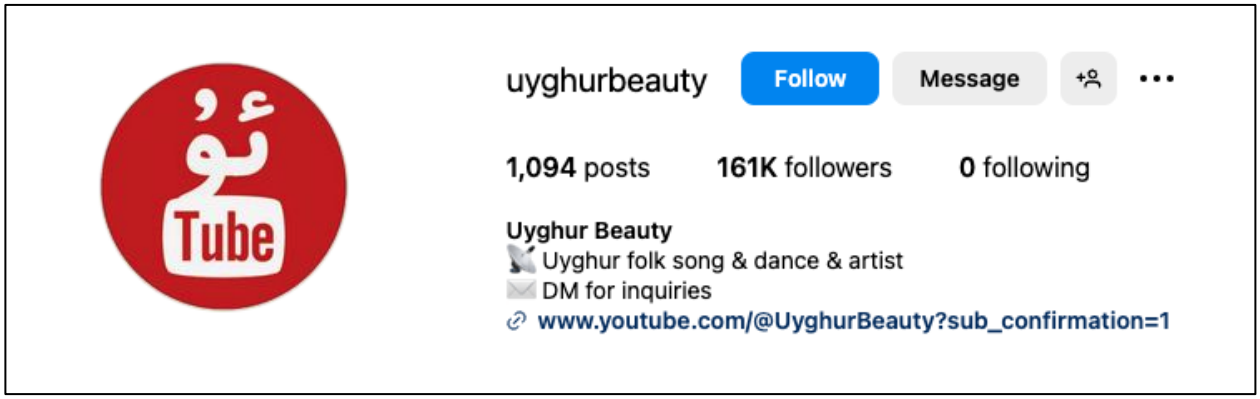


Figure 15: @uyghurbeauty Instagram Profile

Though passive OSINT techniques do not allow for a definitive identification of the channel's operator or its real geographic base of operations, it remains an emblematic example of a Chinese media asset that uses its high-reach and high-frequency of posting to co-opt and dominate social chatter around Uyghur-related keywords and hashtags. This modus operandi has effectively skewed YouTube content to nearly 52.6% pro-China, and provides a qualitative contrast to the official media sources (CGTN, South China Morning Post, etc.) that skewed YouTube content for "Xinjiang" to 52% pro-China.

Another curious YouTube channel we observed is the similarly-named @uyghurgarden (note that @uyghurbeauty repeatedly shares the hashtag #uyghurgarden on Instagram, suggesting though not proving a connection between the two channel operators). @uyghurgarden has 937K

subscribers and nearly 700M net views. The YouTube profile claims it is based in the United States and was opened in 2014, yet at the time of writing the oldest video on the channel dates from just September 2023. Given the fact that archived screenshots of @uyghurgarden from December 2022 depict now-deleted videos,[18] one can conclude that the channel's operators are performing intensive content curation for unclear reasons.

Figure 16: @uyghurgarden YouTube Profile

The account growth witnessed by @uyghurgarden strongly suggests artificial inflation of subscribers. The channel has grown from around 50K followers to August 2022 to nearly 1M by July 2024 as measured by Social Blade, translating into 20X growth over 24 months based on repetitive, relatively low-caliber short videos. Subscribers are up nearly 268% for the past 30 days, while views increased by 282%. In the absence of high-profile viral content, these subscriber and views growth metrics appear highly inorganic and are suggestive of inauthentic amplification.

Figure 17: SocialBlade Data for @uyghurgarden

[18] Web.archive.org. (2022, December 12). Uyghur Garden. YouTube. Retrieved from https://web.archive.org/web/20221212154805/https://www.youtube.com/@uyghurgarden

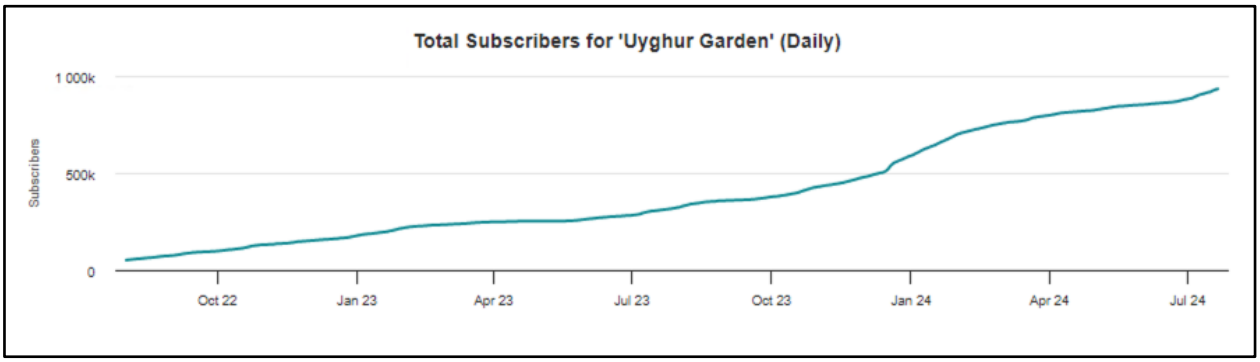


Figure 18: SocialBlade Data for @uyghurgarden, cont'd

Closer analysis of the two accounts suggests that they are indeed likely managed by a single operator. Firstly, the generic wording of @uyghurgarden's account description closely resembles that of @uyghurbeauty's: "Welcome to the Uyghur Garden! Discover Uyghur people and their delicious food." Furthermore, both accounts use similar custom-designed red symbols that further indicate a connection. Lastly and most tellingly, an archived screenshot from December 2022 shows that @uyghurgarden had added @uyghurbeautfy as a featured account to its channel profile.[19] As this step must be undertaken by the channel operator, we can infer with high confidence that the operators of @uyghurgarden and @uyghurbeautfy are either shared or in some degree of contact.

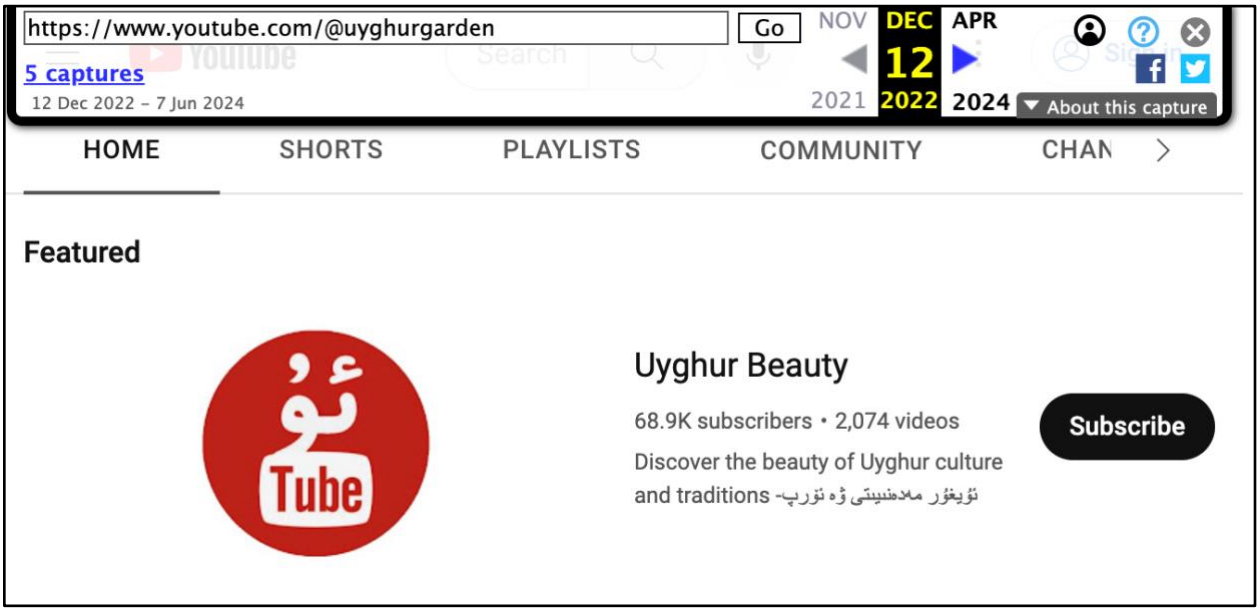


Figure 19: Web Archive Screen Capture of @uyghurbeauty from 12/12/2022

Though open-source techniques are unable to definitively prove a connection between the two channels, both exemplify the well-documented CCP tactic of using idyllic coverage of folk culture to whitewash severe human rights abuses and cultural erasure of China's ethno-religious minorities. They also illustrate the significant resources being allocated to the

[19] Ibid.

cultivation of media assets that are then used to target international audiences by seeding benign content that subsumes substantive content around sensitive domestic issues for the CCP.

APPENDIX II

Supplementary Survey Data

Coefficient plots for survey regression analysis was performed across key questions pertaining to China and human rights. Regressions were performed in a step wise fashion to examine the robustness of main effects and check against heteroskedasticity.

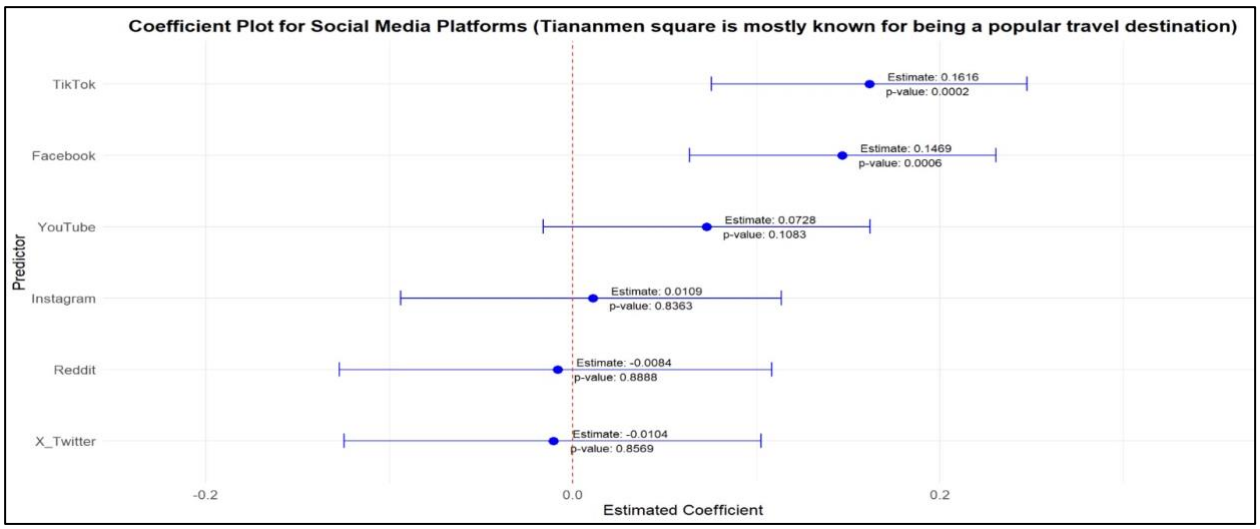


Figure 20: Coefficient Plot for Social Media Platforms (Tiananmen Travel)

Coefficient Plot for Social Media Platforms (Tiananmen Square is mostly known for being a popular travel destination)

This plot illustrates the impact of various social media platforms on the perception that Tiananmen Square is predominantly recognized as a travel destination rather than a historical site of political significance. TikTok emerges as the leading platform, significantly amplifying this perception among its users. In comparison, Facebook[20] also shows a notable effect, albeit less pronounced than TikTok. Other platforms like Instagram, YouTube, Reddit, and X (formerly Twitter) demonstrate comparatively minor impacts.

[20] Here it is important to note that significant associations with Facebook use and favorability towards China's human rights record did not survive more scrupulous pressure testing of our models, but Tik Tok use did uniquely among platforms. See extended modeling/factor testing on favorability towards China's human rights records here: https://docs.google.com/document/d/1wVEtKmk7XsZy1QL5llxHqcRQh9SDDxXSVtURRDnUbXE/edit?usp=sharing

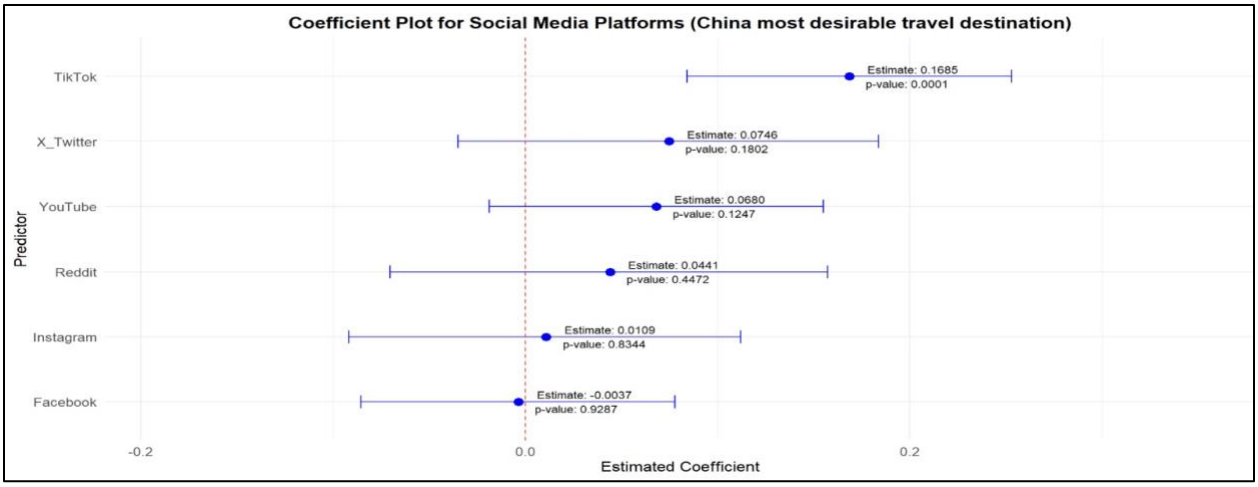


Figure 21: Coefficient Plot for Social Media Platforms (China Travel)

Coefficient Plot for Social Media Platforms (China as the Most Desirable Travel Destination)

This plot examines how different social media platforms influence the view of China as the most desirable travel destination. TikTok once again stands out, significantly enhancing positive perceptions of China as a premier travel spot. Reddit and X show moderate influence, whereas platforms like YouTube, Facebook, and Instagram have minimal or negligible effects in shaping this perception.

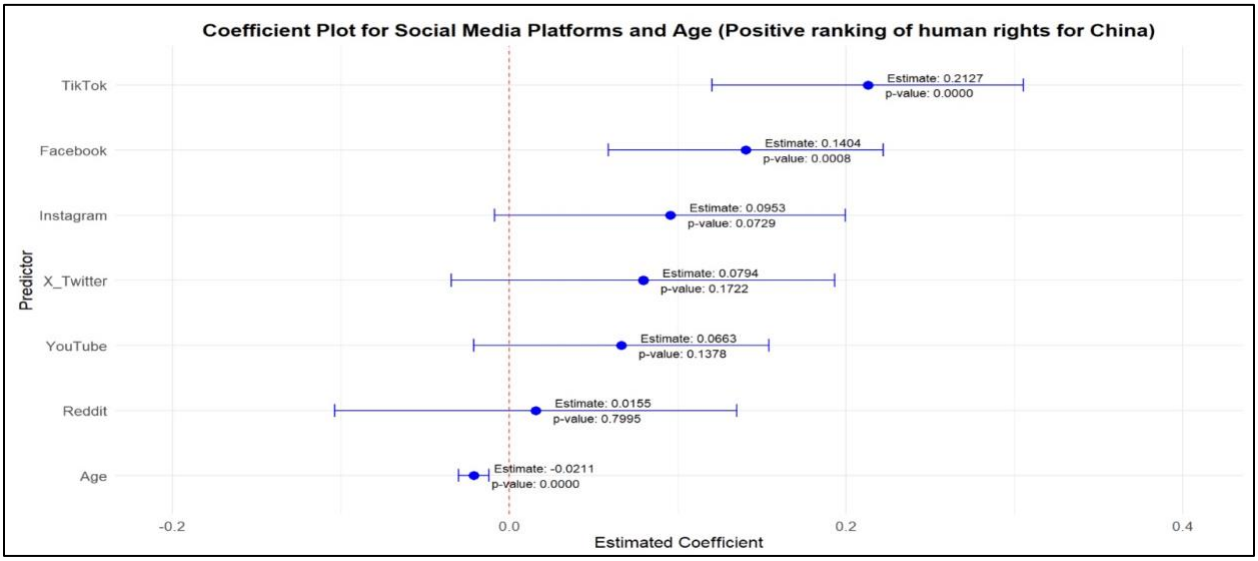


Figure 22: Coefficient Plot for Social Media Platforms (China Human Rights)

Coefficient Plot for Social Media Platforms and Age (Positive Ranking of Human Rights for China)

This figure evaluates the role of social media platforms and user age in positively ranking China's human rights record. TikTok prominently leads, indicating a substantial impact on improving perceptions of China's human rights among its users. Facebook follows with a significant but smaller effect. Instagram, X, and YouTube show modest influences, while Reddit and user age exhibit the least impact on this perception.

Overall, TikTok consistently demonstrates the highest influence across all examined perceptions, highlighting its powerful role in shaping user views on key issues related to China. Facebook also shows significant, albeit lesser, influence compared to TikTok in altering user perceptions.

APPENDIX III

Engagement Metadata (YouTube and Instagram)

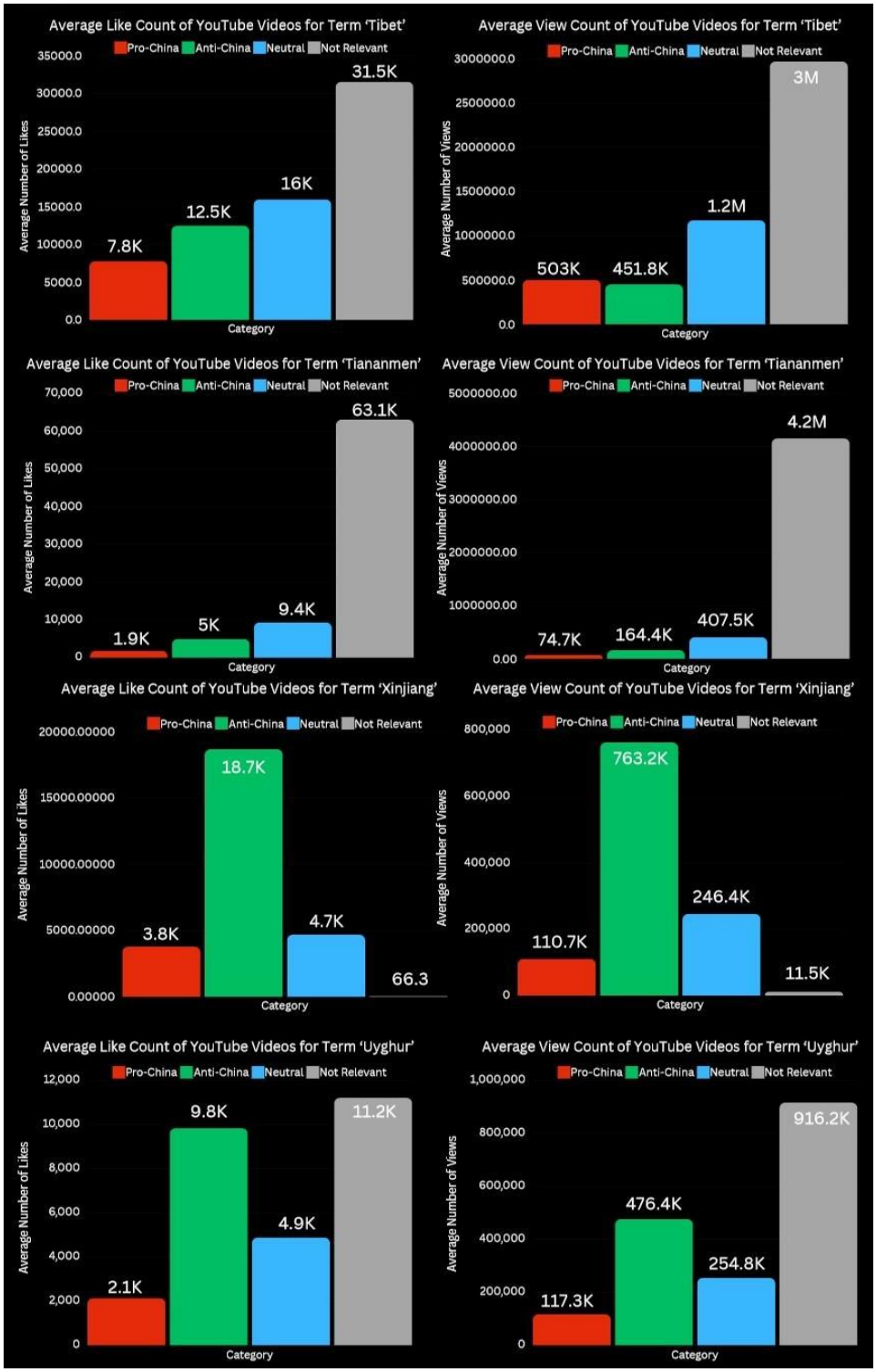


Figure 23: YouTube Engagement Metadata

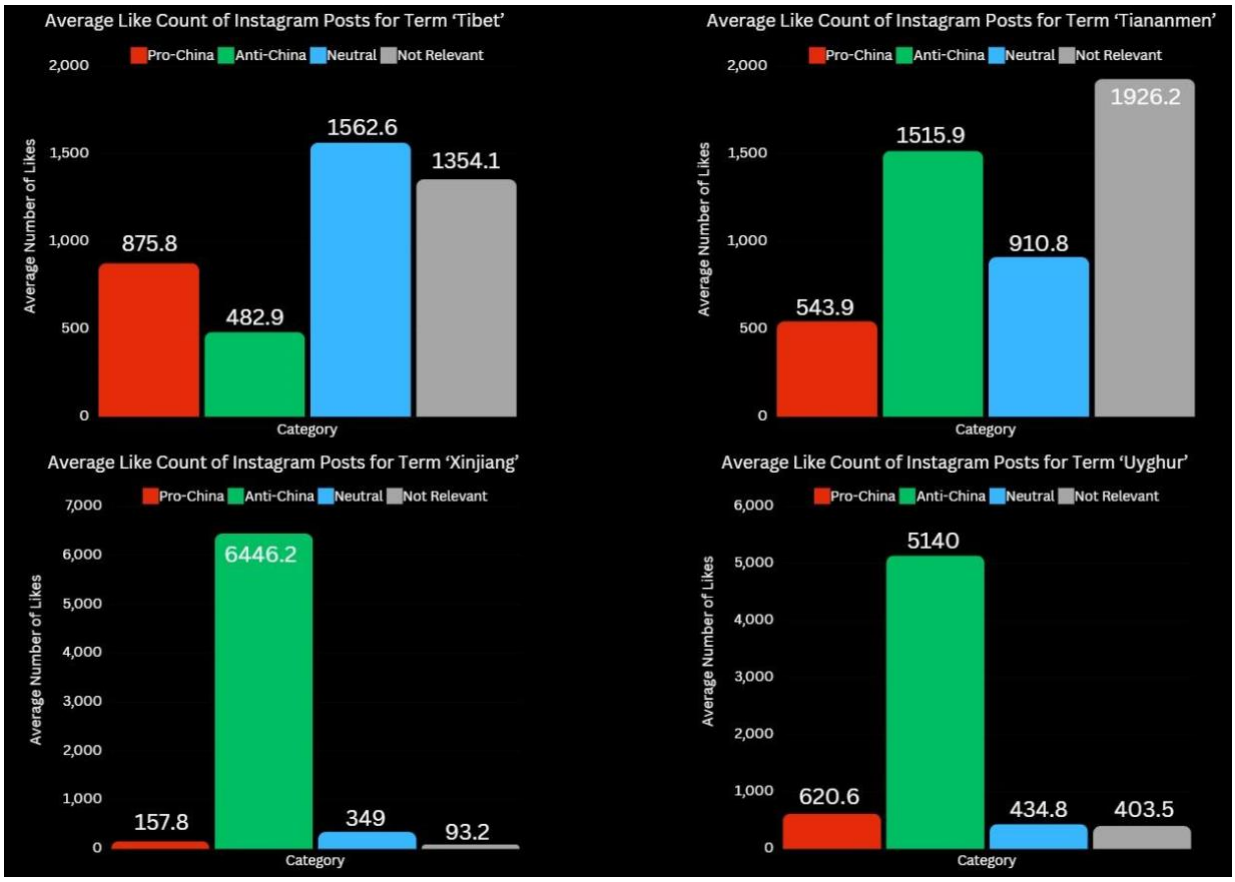


Figure 24: Instagram Engagement Metadata