

Information control by public punishment: The logic of signalling repression in China

Lotus Ruan
University of Toronto

Jeffrey Knockel
University of Toronto

Masashi Crete-Nishihata
University of Toronto

China Information
1–25

© The Author(s) 2020



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0920203X20963010
journals.sagepub.com/home/cin



Abstract

When does repression of online expression lead to public punishment of citizens in China? Chinese social media is heavily censored through a system of intermediary liability in which the government relies on private companies to implement content controls. Outside of this system the Chinese authorities at times utilize public punishment to repress social media users. Under China's regulatory environment, individuals are subject to punishment such as fines and detention for their expressions online. While censorship has become more implicit, authorities have periodically announced cases of repression to the public. To understand when the state escalates from censoring online content to punishing social media users for their online expressions and publicizes the punishment, we collected 468 cases of state repression announced by the authorities between 1 January 2014 and 1 April 2019. We find that the Chinese authorities most frequently publicize persecutions of citizens who posted online expression deemed critical of the government or those that challenged government credibility. These cases show more evidence of the state pushing the responsibility of 'self-regulation' further to average citizens. By making an example of individuals who post prohibited content even in semi-public social media venues, the state signals strength and its determination to maintain authority.

Corresponding author:

Lotus Ruan, Citizen Lab, Munk School of Global Affairs and Public Policy, University of Toronto, 315 Bloor Street West, Toronto, ON M5S 0A7, Canada.

Email: lotus@citizenlab.ca

Keywords

repression, social media, information control, public punishment, online expression, public transcript, self-regulation

Social media companies operating in China are required to censor content according to government regulations.¹ For most users in China, the impact of China's information controls is in the denial of information. However, individuals sometimes come under pressure for content they posted online with consequences ranging from 'invitations to tea' (a euphemism for being questioned by the police), to arrests, detention, and prosecution. While social media censorship is delegated to private companies and the government typically does not publicly claim responsibility for specific instances of censorship, the Chinese authorities have at times taken action to punish social media users for posting certain content and they have publicized these cases of punishment.

The escalation from censorship to legal punishment by the state is presumably because certain content is considered highly sensitive to the state and censorship alone cannot offset the potential threats. What types of content or users constitute threats to the state? Why does the government make these cases known to the public?

In this article, we present a study on the targets and goals of China's information control over social media from the perspective of the state. We analyse when the state takes legal action (i.e. arrests, detention, and prosecution) in response to online expression on WeChat, one of the most popular social media platforms in China. WeChat provides a variety of communication functionalities including instant messaging (e.g. one-to-one chats and group chats), WeChat Moments (a functionality that resembles Facebook's Timeline where users can share text-based updates, upload images, and share short videos or articles with their friends), and the WeChat public account platform (a blogging-like platform that allows individuals as well as businesses to reach out to general audiences). We focus specifically on cases of arrest, detention, and prosecution related to the use of social media that are announced by the authorities in public venues such as the official websites or social media accounts of police units.

Our focus on state-highlighted cases is inspired by the literature on 'public transcripts',² which emphasizes the importance of public interaction between citizens and political authorities in everyday settings. Whereas censorship aims at concealing information from citizens, public script in authoritarian regimes, such as the public display of punishment, is an explicit propaganda attempt to shape the political and social values of the citizenry.³ Under the current Xi Jinping administration, China has seen an increasing number of public displays of punishment and forced public confessions.⁴ Analysing official announcements of punishment helps us better understand the targets and goals of China's information control, where the selection of what and whom to punish often seems arbitrary and stems from political considerations.

Based on an original dataset of 468 official reports of punishment of WeChat users, we find that most of the cases were in reaction to a user's post in group chats (52.6 per cent) and WeChat Moments (28.2 per cent). Overall, we find that, among a variety of content that led to state repression, authorities most frequently highlighted punishment of individuals who shared information that had not been verified by the authorities (31.4 per cent) or individuals who criticized the government (29 per cent).

Dismissed by the authorities as ‘rumour’, this type of content covers a wide range of topics including information surrounding violent cases such as child abduction, issues of public health, casualties in accidents, and government policies. Meanwhile, criticism of government ranges from criticism against rank-and-file police, government or party leadership, official party ideology and state policy, to threats of individual actions against the government.

Overall, our results suggest that the Chinese state invokes its repression apparatus to control information that is critical of the government or information that challenges government authorities and credibility. In addition to the deterrence effect of making an example of individuals who post undesirable content, the state is also signalling its strength and determination to maintain its authority.

This article proceeds as follows. We first review the literature on the theory and practice of information control in China including censorship, repression, and propaganda. We then describe our method for identifying cases of repression related to users’ use of WeChat. We conclude with discussions of the implications of our results and how our study can inform research on the repression of social media use in China.

Literature review

In this section, we provide an overview of China’s regulatory environment for online content, theories on China’s information control system including its censorship and propaganda strategies, and the recent trends of crackdowns on social media users under President Xi Jinping.

Theories on China’s information control system

Information controls are instrumental to the rule of the Chinese Communist Party (CCP).⁵ A large body of literature studies the mechanisms, targets, and goals of China’s online censorship regime. Yet, the types of content that the CCP is targeting remains a topic of debate. Gary King, Jennifer Pan, and Margaret Roberts argue that the Chinese authorities target online content that has ‘collective action potential’ but that they tolerate government criticism.⁶ Recent studies present counterpoints to this theory and show a more nuanced picture. Studies of keyword-based censorship on chat applications,⁷ live streaming applications,⁸ and mobile games⁹ found that censored keywords included references to collective action and criticism of the government. An analysis of leaked logs from the censorship department of SinaWeibo showed that content related to collective action, political humour, and government criticism is censored at a similar rate.¹⁰ Findings from other studies revealed that SinaWeibo censored CCP-related keywords for a longer period than those referencing opposition or protests, leading to the conclusion that the goal of the CCP’s censorship practice is to protect its one-party rule.¹¹

There are at least two reasons why observations of social media controls described by King, Pan, and Roberts¹² are at variance to other research.¹³ First, China’s social media censorship is operated through a system of intermediary liability, which allows the government to push responsibility for information control to the private sector.¹⁴ In 2017, Chinese regulators pushed information control even further, making individual users responsible for the content they post online.¹⁵ This decentralized system has created a

Table 1. Examples of domain-specific laws pertaining to content regulation.

Date in effect	Domain	Regulation
7 August 2014	Instant messaging services	Interim Provisions on the Administration of the Development of Public Information Services Provided Through Instant Messaging Tool
1 June 2017	All Internet services	Cybersecurity Law
1 October 2017	Online forum	Internet Forum Service Management Regulation
1 October 2017	Online comments	Internet Thread Comments Service Management Regulation
8 October 2017	Online chat groups	Management Rules of Internet Group Information Services

Source: 政策法规 (Policies and laws), 中华人民共和国国家互联网信息办公室 (Cyberspace Administration of China), http://www.cac.gov.cn/zcfg/fl/A090901/index_1.htm, accessed 30 June 2020.

principal–agent problem¹⁶ and makes it difficult to conclude whether censorship decisions come from the state (principal) or companies (agent).¹⁷ Companies may defy censorship directives out of their own business interests¹⁸ or apply a broad censorship strategy to avoid official reprimand.¹⁹ Second, while there are various laws and regulations pertaining to content controls (see Table 1 for examples of domain-specific Internet regulations), few offer precise definitions of prohibited topics. The vaguely defined regulations encourage companies to over-censor²⁰ and allow for arbitrary targeting.²¹

Trends of social media controls in China

In recent years, social media censorship has become more subtle and fine-grained in the way that different kinds of users are targeted.²² Skilful authoritarian regimes like China control social media strategically to their advantage by manipulating online public opinion,²³ showing responsiveness to disgruntled citizens,²⁴ and keeping corruption at local levels in check.²⁵ Yet, the state has not reduced the use of its repression apparatus, which has also become more targeted, pre-emptive, and with a greater focus on ‘psychological coercion’.²⁶

The majority of social media users in China experience information control via censorship.²⁷ For some outspoken individuals, however, the state’s repressive tactics could be more explicit. They may face repercussions ranging from arrest to prosecution. The state has propagated some of these cases of crackdowns.²⁸ One of the most notable waves of publicized crackdowns came amid the Chinese state’s campaigns against ‘online rumours’ which have taken place since the early 2010s.²⁹ Existing literature has delved into the problematic use of the term rumour (or ‘fake news’, ‘misinformation’ in other contexts) by the state and those in power.³⁰ While some stress the ‘false, unverified, and defamatory nature of rumour’,³¹ rumours also act as a form of social protest or ‘weapons of the weak’³² for average people to challenge official narratives. Recent research demonstrates that rumours decrease citizens’ trust in the government and support of the regime, regardless of their background.³³ Authorities often abuse or demonize the term rumour to crack down on information deemed threatening to the state³⁴ and to produce a deterrent effect on the public.³⁵

Under the Xi Jinping administration the campaign against rumours has accelerated.³⁶ In 2016, the Chinese authorities stated that content on social media can be collected and used as ‘electronic data’ to investigate legal cases.³⁷ Besides pursuing criminal charges, police have the power to hold someone in ‘administrative detention’ (i.e. the arrest and detention of individuals by the police without proceeding further to trial) for ‘spreading rumours’.³⁸ Administrative detention has been criticized because of its potential for abuse by China’s police.³⁹ Jason Q. Ng’s study of censorship on WeChat confirms that rumour can be an ambiguous term to censor not only falsehoods but also harmless content.⁴⁰

While existing regulations seem to suggest that the state’s information control apparatus is geared towards controlling public opinion and preventing collective action, the criteria for interventions and punishment by the authorities are left unclear.⁴¹ Mary Gallagher and Blake Miller analyse the types of SinaWeibo users being reported back to China’s security apparatus and argue that rather than focusing exclusively on restricting sensitive content, the state permits open discussion among ordinary people while keeping close tabs on influential non-party ‘thought leaders’.⁴²

This line of analysis (i.e. focusing on the question of ‘who’ rather than ‘what’) offers a new perspective in understanding the logic of China’s information controls. Yet, research on state repression of online space, especially with regard to newer platforms such as WeChat, is relatively limited. Unlike SinaWeibo where it is easy to track how influential a user or an article is, WeChat does not display how many friends a user has, how many times a message is forwarded or viewed (except for articles published on the WeChat public account platform). However, previous research comparing the potentials of WeChat’s different functionalities in facilitating public opinion suggests that its public account platform allows for the widest dissemination of information whereas group chats and Moments target mostly close-knit social circles.⁴³ Although we do not have accurate metrics of how influential a certain user or content is, analysing the frequency of each WeChat functionality targeted by the state offers insight into motivations behind state controls.

Rationale for focusing on cases highlighted by the state

To understand the mechanics of information control, we should observe the system of controls from the state’s perspective and analyse the types of content and users punished by the state itself. Inspired by James Scott’s theory of public transcript and existing work on China’s propaganda regime, we focus on cases self-reported and propagated by the authorities.⁴⁴ A great amount of research has been dedicated to understanding public transcript in authoritarian regimes, and findings show that the aim of public transcript is to ‘reinforce [o]ffence categories’⁴⁵ and to confirm the importance of ideological conformity and the risks of non-conformity.⁴⁶ Rachel Stern and Jonathan Hassid demonstrate how the atmosphere of uncertainty, coupled with occasional crackdowns, lead to self-censorship among the most outspoken professionals in China.⁴⁷

Public transcript is a staged public performance through which political authorities hope to convey to the general public what behaviours are unacceptable; it is also a public performance through which ordinary people adopt the forms of deterrence and respect

for power holders that are needed to avoid punishment.⁴⁸ ‘Pedagogy through prosecution’ is effective in generating fear and reining in opposition while minimizing the amount of actual prosecution.⁴⁹ Public transcript is therefore a form of propaganda, which, as Damien Ma and Neil Thomas point out, is instrumental to China’s political system.⁵⁰ While censorship is aimed at concealing information from citizens, propaganda (e.g. publicizing punishment of an individual or a type of behaviour) is an explicit attempt to inject political and social values into citizenry.

Existing work theorizes that the goals of the CCP’s propaganda include instilling nationalism into the public,⁵¹ manipulating public opinion,⁵² maintaining the state’s agenda-setting power, and sustaining regime legitimacy.⁵³ While it is questionable how effective China’s propaganda strategies are in the digital age,⁵⁴ Haifeng Huang argues that the value of propaganda in authoritarian states lies not in persuasion but in ‘signalling’⁵⁵: by producing a high level of propaganda, governments show their citizens who is in charge, and in doing so, they discourage potential rebellions.

The rationale of this study is further justified by that fact that public displays of punishment and forced public confessions have become regular fare under Xi Jinping’s administration.⁵⁶ Similarly, China’s Internet controls and security apparatus have begun publishing ‘typical Internet crimes’ on official websites and social media accounts in an attempt to ‘educate Internet users about the law’.⁵⁷ Paying attention to cases of punishment that the authorities have handled is helpful for understanding the logic of information controls in authoritarian states.

Methodology

This study assembles an original dataset of 468 unique cases of state repression self-reported by the authorities between 1 January 2014 and 1 April 2019. Reports of these cases were published in highly visible official channels accessible to the public. We analyse cases of repression involving content users posted on WeChat. WeChat has three major social functionalities where users can post content: (1) chats, which includes one-to-one chats and group chats; (2) WeChat Moments; and (3) the WeChat public account platform.

We gathered data through searching three well-known sources. We only considered cases that (1) involved the use of WeChat; (2) documented state response in reaction to such use; and (3) were made public by the authorities themselves.

The first source of data consists of social media posts from 263 official SinaWeibo accounts of Chinese Internet police units. In 2015, China’s Ministry of Public Security launched a campaign to ‘normalize Internet inspection work’. Police units across China launched social media accounts to publicize their work policing ‘typical Internet crimes’.⁵⁸ Since the SinaWeibo accounts of police units uniformly include the location of the police unit and the phrase ‘Internet Police Patrolling and Enforcing Law’ (网警巡查执法) in their account name,⁵⁹ we used this phrase and performed an account search on SinaWeibo and identified 263 verified official accounts of Internet police units. These accounts covered different levels of police units from 31 provincial-level regions in mainland China (i.e. 22 provinces, five autonomous regions, and four direct-administered municipalities). Most accounts followed the official naming patterns while a handful used alternative phrases with the same meaning (e.g. 深圳网警,

Shenzhen Internet Police). We scraped all original posts from the 263 official accounts. This source generated 167 unique cases in total.

The second source of data comes from the Cyberspace Administration of China (CAC). Founded in 2014, the CAC is the top Internet regulatory authority at the central level and it enjoys broad authority such as issuing orders to the Ministry of Public Security and its Internet police units to regulate Internet content and combat online crime.⁶⁰ As a national organ, the CAC centralizes all cases where a person is arrested for posting ‘harmful information’, and it publicizes select ‘typical cases’ on its website. We wrote a script to search and collect all public reports from the CAC as of 1 April 2019 with the keywords ‘WeChat’ and ‘detention’. Removing duplicates, this source generated 187 cases in total.

The third source of data comes from reports from WeChat public accounts. Although Internet police units also have official WeChat public accounts, it is difficult to scrape these accounts and all of their previous posts because WeChat only allows searching and viewing of a public account’s historical articles on the mobile or desktop version of the application and because the URL of each article is difficult to predict. Due to the scope and timeline of this project, we resorted to Sogou, WeChat’s official partner search engine. We searched for reports that mentioned WeChat and detention. After removing duplicates, this source gave us a total of 155 unique cases. These search results likely have been filtered by WeChat and/or Sogou due to censorship compliance. However, since our objective is to understand the types of content or users that Chinese authorities target and wish to communicate to the general public, such censorship bias actually helps with our research.

We then manually read each of these cases and removed duplicates across the three data sources. In the end, we gathered 468 unique cases of state repression. Each case of state repression is grouped into content categories based on a code book we developed for this study (see Table 2). Content categories reference the topic category of content or behaviour targeted by the state’s security apparatus. One co-author (a fluent Chinese speaker) manually coded the data. Another performed inter-rater reliability checks on a randomized sample of the coded cases to ensure consistency.

Findings

In this section, we present an overview of our findings. Our data consist of cases spanning from 2014 to 2019. Overall, we find that the authorities featured punishment of undesirable content in WeChat’s group chats most frequently during this period. The undesirable content detailed in official reports was predominantly about government criticism and information labelled as rumours by the authorities. Most of these reports (184 out of 468) were published between January and April 2019, followed by 2012 which saw a total of 102 cases published over the course of 12 months.

WeChat functionality analysis

Most police and state media reports (435 out of 468) mentioned the WeChat functionality that was used to post the undesirable content. Figure 1 shows the distribution of WeChat

Table 2. Breakdown of all categories used for categorizing content.

Content category	Definition
Government criticism: local police	Includes criticism against rank-and-file local police
Government criticism: government/party leadership	Includes criticism against a government or party leader
Government criticism: party ideology and state policy	Criticism against official party ideology or state policy
Government criticism: threats of individual action	Mentions actions against governments organized by an individual (as opposed to collective action)
Collective action	Mentions protest, petition, and so on
Rumour: public health	Mentions uncertain, unverified information or partial truth of stories related to public health, such as vaccines and viruses
Rumour: violence	Mentions uncertain, unverified information or partial truth of stories related to murder, robbery, and other violent crimes
Rumour: casualty	Mentions uncertain, unverified information or partial truth of stories related to casualties caused by an accident
Rumour: reputation	Mentions uncertain, unverified information or partial truth of stories that aim at or result in the damage of the reputation of an individual, an organization, or a place
Rumour: disaster	Mentions uncertain, unverified information or partial truth of stories related to a man-made or natural disaster
Rumour: government policy	Mentions uncertain, unverified information or partial truth of stories related to a government policy
Rumour: context unclear	Mentions a person arrested for spreading rumours without providing any information on the content of the alleged rumour
Terrorism	Mentions terrorism, global or domestic
Government policy	Mentions government policy in a neutral way
Moral offence	Mentions the violation of social norms or other forms of social deviances that do not directly charge political authorities such as drug use, sales of illicit goods, gambling, and prurient interests
Ethnic groups	Mentions ethnic groups and related disputes
Personal	Mentions a person other than government or party officials, or if content does not specify who the person is
Party discipline	Mentions punishment of officials and party members involved in corrupt or other disciplinary behaviour such as embezzlement
Lack of nationalism	Invokes irreverence toward symbols or deceased role models of China or the party; state action may be justified on the grounds of nationalism
Miscellaneous	The content topic is unclear from the text of the report, for example if a report says someone is arrested for spreading 'harmful information' without specifying the content

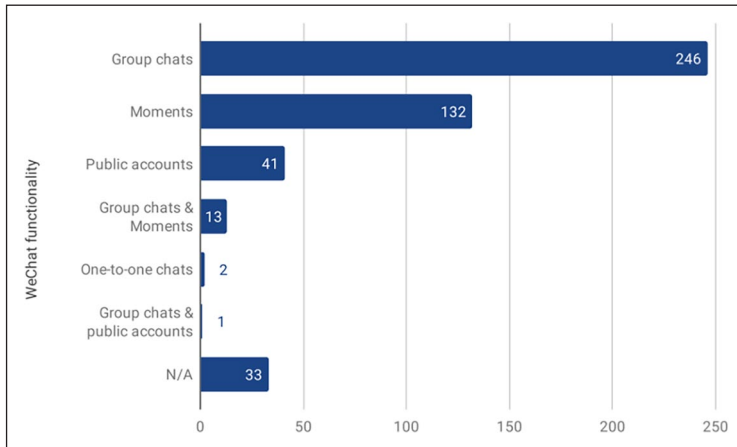


Figure 1. Distribution of cases by functionality, 1 January 2014–1 April 2019.

Sources: Compiled from authors' dataset comprising 263 official SinaWeibo accounts of Chinese Internet police units; cases of Internet-related detention published on the CAC website; and official reports mentioning 'WeChat' and 'detention' on WeChat public accounts.

functionalities mentioned in the dataset. The majority of police reports highlighting prohibited behaviours in WeChat was found in functionalities with the strongest social component. Over half of the cases (52.6 per cent) referenced content shared in group chats, followed by those referencing content posted on WeChat Moments (28.2 per cent). Forty-one reports highlighted the punishment of users who posted undesirable content on WeChat's public account platform.

Only two reports on the punishment of WeChat users were about their use of one-to-one chats. One of them was an employee at a local environmental protection bureau who accepted a bribe in the form of 'WeChat red packet' in a one-to-one chat with someone who violated environmental regulations. The other case involved a WeChat user detained by the police for insulting other people in a one-to-one chat. Among the 468 reports, 33 did not detail where in WeChat the questionable content was posted. Fourteen reports mentioned more than one WeChat functionality without specifying which one directly led to the punishment.

The pattern of the state targeting functionalities that affect larger audiences is consistent with previous studies,⁶¹ with some nuances worth further exploring. By design, WeChat's public account platform is most effective in disseminating information to the largest audience since the content can reach anyone including those that are not WeChat users. However, in our dataset, prohibited behaviour on WeChat's public accounts was not highlighted as often in official police reports as those in group chats or WeChat Moments, where content is only available to a user's contacts.

A possible explanation can be drawn from studies on mass communication and formation of public opinion.⁶² While there are limitations on their ability to form public opinion or mobilize collective action, WeChat's group chats and Moments are based on close-knit networks where social interaction features strengthen social bonds among

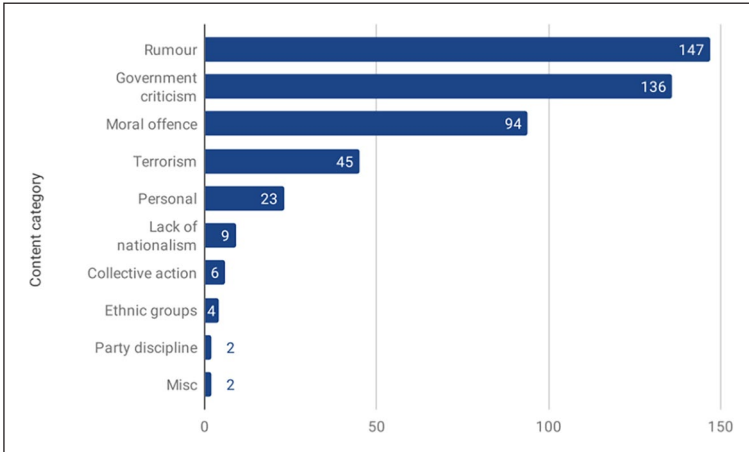


Figure 2. Distribution of content by category, 1 January 2014–1 April 2019.

Source: Authors' dataset.

people. Real-time chat is potentially a powerful venue for people to incubate opinions more privately in small groups and to potentially address coordination problems.

Content analysis

To better understand each case of state repression, we analysed the content detailed in each report, grouping them into 11 primary content categories based on an interpretation of the underlying context of each case. We found a diversity of targeted content. In addition to frequently targeted content referencing collective action and government criticism,⁶³ we also found politically charged content such as messages related to nationalism (or, rather, lack thereof) and non-political content such as references to illicit goods and gambling.

Of the 468 reports, the majority of individuals who faced state repression in the form of fines and/or detention were those who criticized or ridiculed an organ or officer of government and those who posted rumours (i.e. information not verified by the authorities or information that represents only a partial truth of an event). The third largest category is content referencing moral offences and non-political crimes, including prurient interests, gambling, and sales of illicit goods. Figure 2 shows the distribution of the types of content targeted by the state's security apparatus by category. Here we describe and provide examples of these categories.

Content referencing 'rumours'

Over 34.1 per cent of cases highlighted by the authorities involved content containing information not verified or acknowledged by official sources. While most cases offered the full context or official explanation of the problematic content, a few reports mentioned only vague information on why an individual was punished. To gain a more

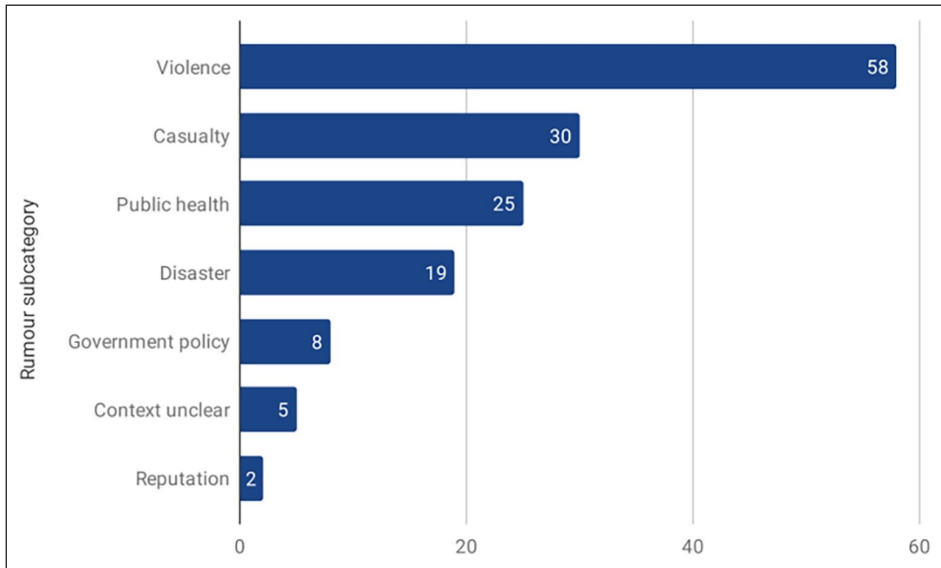


Figure 3. Distribution of subcategories of rumour-related content, 1 January 2014–1 April 2019. Source: Authors' dataset.

nanced understanding of the types of rumours highlighted by the authorities, we manually classified the 147 rumour-related messages into seven subcategories based on their context and relevant events. Figure 3 shows the distribution of rumour by subcategory.

We found that the majority of cases under this category (39.4 per cent) concerned spreading information about uncertain, unverified accounts, or partial truth of a violent event including stories related to murder, assault, robbery, child abduction, and other violent crimes targeting average citizens. Rumours that mentioned unverified information about public health-related issues such as an outbreak of Ebola, H7N9, HIV/AIDS, and those that mentioned a natural or man-made disaster such as earthquakes were targeted at a similar rate (17 per cent and 12.9 per cent respectively).

Over 20 per cent of the rumours mentioned an inaccurate number of casualties in a natural or man-made disaster. In most cases, these rumours were based on real events that made national news or on local accidents that occurred where the information sharers were based. Rumours were posted soon after the accident had happened when information about the cause, casualty, and loss was not confirmed by the authorities.

For example, a WeChat user surnamed Chen in Zhenjiang City, Jiangsu Province, was detained for three days by the local police for exaggerating the number of deaths and injuries in the 2015 Tianjin explosions on WeChat Moments. The 2015 Tianjin explosions refer to a series of explosions on 12 August 2015 that killed 173 people and injured hundreds of others at a container storage station at the port of Tianjin. The police report details the full content of Chen's post on WeChat Moments: 'The Tianjin explosion has caused the death of 482 average citizens. Fifty-two people were heavily injured and are still in critical condition. Four teams of fire-fighters died. The explosion affected areas within a 1-km radius. Residents in a nearby community were all dead from the accident.'

According to the police report, Chen deleted his post within half an hour after posting. However, he still faced punishment because ‘his post has caused a pernicious influence’ (a vague phrase commonly used in police reports). Past research shows that the 2015 Tianjin explosion incident was one of the most frequently censored events of that year⁶⁴ and that online discussion was carefully managed.⁶⁵

In another report by the police department in Xiantao City, Hubei Province, a WeChat user surnamed Wu was detained for five days for ‘fabricating facts and disrupting public order’. On 12 January 2019 Wu was reported to have filmed and posted a 12-second video of a local car crash in a WeChat group in which he said, ‘There’s been a big accident. A few people died. There’s a big accident between Xiantao and Magang.’ According to the police report, while there were a series of car crashes that day, there were no casualties. Wu was under police detention a day after he posted the video.

Eight rumour-related cases included references to unverified or inaccurate information about government policies, particularly those related to people’s livelihood. A user in Wuqiang, a county in Hebei Province, was detained for spreading in a group chat inaccurate information about a government decree that aimed at replacing coal with cleaner-burning natural gas. Similarly, a Yunnan-based public account published a post about government plans to build a railway in Yunnan. The account was permanently banned because the post was inaccurate and that there was no such construction plan.

Two out of the 147 rumour-related cases included misinformation about an event that was otherwise harmless except that it damaged the reputation of a locality. Both were posted in public accounts whose target audience comprised mainly residents of that locality. For example, a user operating a WeChat public account in Hai’an City, Jiangsu Province, was detained for two days and asked to ‘apologize to fellow netizens’ for ‘fabricating facts and disrupting public order’. According to the police report, the user posted a video of a theft filed in another city in the public account he was running and claimed that the theft happened in Hai’an.

Authorities did not always release details on what types of rumour an individual allegedly spread. About 3 per cent of rumour-related cases included only the punishment of rumour-mongers without specifying the content of the alleged rumours. In August 2017, for instance, a 58-year-old residing in Xinjiang was subject to administrative detention for ‘spreading a large number of political rumours’ in WeChat. In another case highlighted by an Inner Mongolia police department, a user served a 10-day administrative detention for ‘making and spreading three false audio-video messages’ in WeChat. The vagueness in these official reports bears resemblance to the ambiguity in rumour-related censorship.⁶⁶

Content referencing government criticism

The second most frequently highlighted cases involved content critical of government. A total of 136 out of 468 cases included critical references to an organ of the government, government leaders and party cadres, official ideologies and state policies, or made threats against the government. To understand whether the authorities were clamping down on a particular type of government criticism, we coded the 136 cases into four subcategories based on the context and target of criticism mentioned in each report. Figure 4 shows the distribution of government criticism-related cases by subcategory.

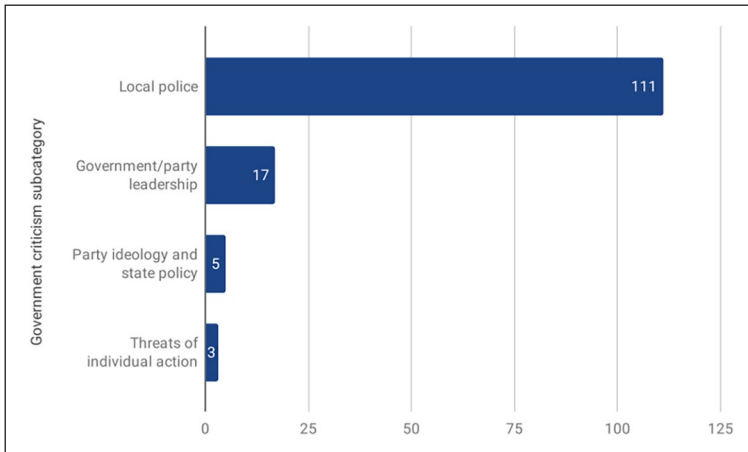


Figure 4. Distribution of subcategories of government criticism-related content, 1 January 2014–1 April 2019.

Source: Authors' dataset.

Interestingly, we found that authorities most frequently highlighted cases involving criticism and complaints against rank-and-file police such as local traffic police (81.6 per cent). Existing censorship studies show that social media platforms primarily control online discussions of high-ranking government officials and party leaders⁶⁷ while tolerating criticism of local officials because the central government supposedly uses those posts to monitor local corruption.⁶⁸ We therefore expected to see more repression targeting social media users who discussed high-level officials rather than the local police force. For example, according to a report by a public security bureau in South China's Guangxi, a WeChat user surnamed Pan was detained for five days for 'insulting the police' in December 2018. Pan reportedly posted photos of police patrolling on his WeChat Moments and captioned, 'Looks like these bandits are feeling very cold' (冷得这些大土匪). Among the 111 cases, four users who 'insulted the police' in group chats faced criminal detention, while the majority faced administrative detention.

Only 12.5 per cent of government criticism cases detailed the punishment of users who criticized a specific government leader or party cadre. Among these cases, only punishment of online expression against government officials at the local level was reported and there was no reference to central government figures. For instance, a public security bureau in Chengdu, Sichuan Province, highlighted a case where a user surnamed Liu was detained for three days for 'damaging other individuals' reputations'. According to the report, Liu filmed a conflict between passengers on a train in January 2019 and posted the video in a WeChat group chat. Liu claimed that two high-level officials of the China Railway Chengdu Group had taken advantage of their status and forced regular train staff to give up their seats. Liu's video circulated widely on the Internet, but authorities dismissed Liu's claims and said that the two seats were designated to the two officials.

A small number of reports mentioned the punishment of individuals who criticized official party ideology and state policy (4 per cent) or made threats against the government (2

per cent). For example, a Xinjiang police department highlighted the punishment of two individuals who ‘maliciously attacked the autonomous region’s stability maintenance policy’ in January 2017. In January 2019, a township-level police department in Guangdong Province published a case involving a WeChat user who allegedly threatened to ‘bomb the public security bureau’. Among all these cases, 75 per cent concluded with administrative detention of various lengths of time.

Content referencing collective action

We found six reports of punishment related to expressions of collective action. In one report, police in Wenzhou City, Zhejiang Province, highlighted a case where a local party member surnamed Yang was sacked for posting inflammatory messages in a WeChat group chat. Yang reportedly encouraged a group of parents to ‘set up a parent committee, draft plans, raise funds, and organize teams to surround the (Wenzhou) education bureau’. Another case involved an online petition in Qianjiang City of Hubei Province in June 2016. According to the report, some party members who founded local chat groups and failed to discourage group members from circulating petition letters or organizing protests were punished.

Content referencing nationalism

We found 10 reports in our database pertaining to individuals who criticized nationalistic sentiments. The most recent case referenced a WeChat user who used derogatory language to describe a firefighter in a group chat. The user called the firefighter ‘dog bear’ (狗熊). The firefighter was acknowledged as a ‘martyr’ by local authorities for his sacrifice in a rescue mission. Notably, we found that the number of cases pertaining to the lack of nationalism highlighted by the authorities in the first four months of 2019 alone was equal to the total number of cases reported between 2014 and 2018.

Content referencing moral offence and personal insults

Not all reports were directly relevant to politics. Authorities also punished those who violated social norms or made reference to morally reprehensible crimes. A total of 93 reports were related to the following categories deemed illegal in China: the sale of illicit goods and services such as drugs, gambling, sexual content or conduct. Another 22 cases involved the punishment of individuals for insulting others (not including insults of government officials) on WeChat.

Relationship between content type and WeChat functionality

We analysed the distribution of targeted content by category in each WeChat functionality. Forty-seven reports did not specify which functionality was used to post the problematic content. We excluded these cases and examined the remaining 421 reports.

Figure 5 shows the distribution of each reported content category across WeChat one-to-one chats, group chats, Moments, and public accounts. At a glance, the majority of

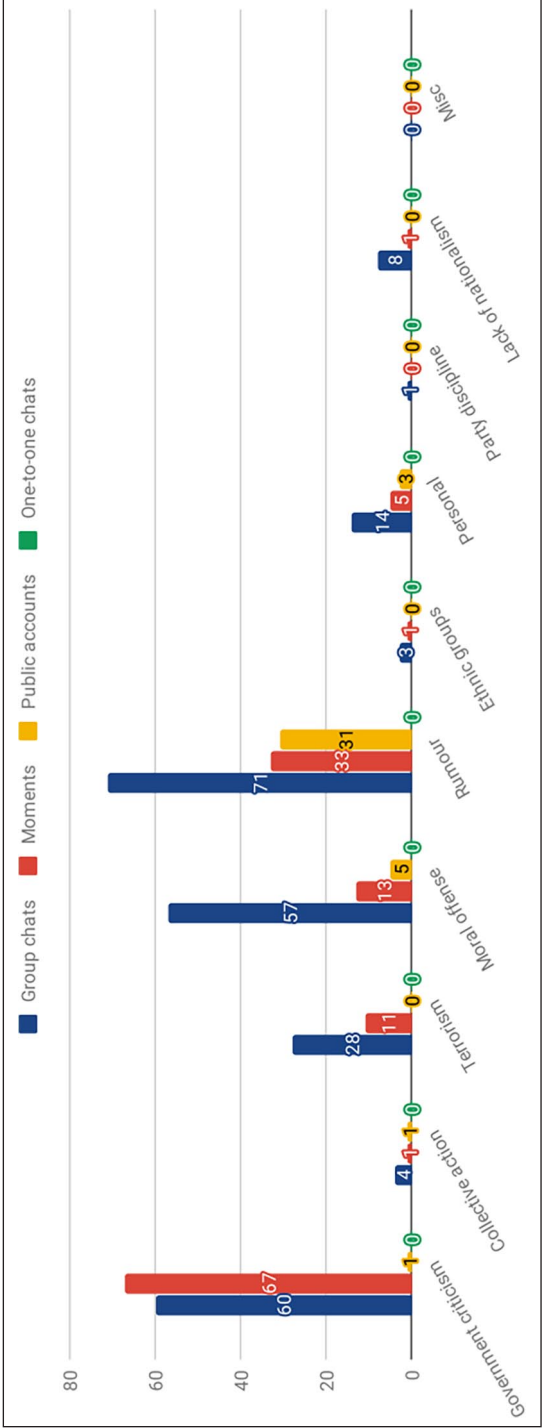


Figure 5. Distribution of content categories in each functionality, 1 January 2014–1 April 2019. Source: Authors' dataset.

content categories receive most of their reports on group chats, fewer on Moments, and even fewer on public accounts and one-to-one messages. To test whether this pattern generally holds true for each content category and that the reports for each content category fall into each of the four WeChat functionalities according to the same distribution, we performed a chi-squared test.

Based on the result obtained ($p < 0.001$), we reject this hypothesis. Rumours were disproportionately reported on public account posts compared to other content categories. Moreover, government criticism was disproportionately reported on Moments posts compared to other content categories. It is unclear whether these results are the product of how government officials choose to report content or on which functionalities users choose to share their posts.

Discussion

What the Chinese state considers unacceptable expressions is reflected in the cases the authorities self-report and propagate to the public. Our study shows that the authorities publicly targeted government criticism and rumours most frequently. The state punished citizens who posted these types of content even when they were shared in semi-public online spaces. Previous studies on Chinese social media censorship suggest that China's information control regime targets criticism and sometimes what appears to be mere discussions of the state.⁶⁹ While these studies use censorship by private companies as a proxy to gauge what the state wishes to suppress, our data mined from official sources confirm their findings as well.

Our data also show that the Chinese government has intensified content controls in online spaces by pushing the burden of censorship further down to the level of individuals. We saw more reports of punishment published by the authorities in the first four months of 2019 than in any single year between 2014 and 2018. The system of self-regulation where the burden of censorship is pushed on to private companies is largely effective,⁷⁰ but there is also evidence that shows that private companies sometimes defy government directives out of their own business interests.⁷¹ It is possible that having realized the problem, the state sees the need to expand its information control apparatus on influential social media platforms. Propagating information on the punishment of individual users who post undesirable content on social media is the state's reminder to its citizenry that ordinary people also bear the responsibility of self-regulation (i.e. self-censorship).

We interpret our finding that the Chinese state most frequently targeted online rumours and government criticism with some caveats in mind. On the one hand, it is possible that the distribution of the highlighted cases reflects the distribution of all cases of state repression including those that are not publicized or highlighted. The punishment of government critics by the authorities is consistent with the patterns of information control demonstrated by existing studies.⁷² The targeting of rumours and rumour-mongers – whether political or not – reflects a continuation and possibly an intensification of China's 'anti-rumour campaign' since the 2010s.⁷³ On the other hand, it is possible that the authorities selectively highlighted these cases while downplaying other more sensitive cases. In August 2017, authorities in Yunnan Province issued a police report

detailing an arrest of a WeChat user who mocked President Xi Jinping in a group chat. However, the authorities soon deleted this report from all its official channels, and discussions on the case were censored on Chinese social media.⁷⁴ Publicizing these cases may invite discussion of the top CCP echelon, and fuel public opinion beyond the government's control. This may explain why there were not many cases of repression against collective action, which King, Pan, and Roberts argued is a main target of China's information controls,⁷⁵ or discussions of high-level government leaders, which previous studies found to be a frequent target of censorship.⁷⁶

Regardless, it is clear from our data that the Chinese state has a low tolerance for content that is critical of the government or that challenges government credibility. There are several explanations and implications behind the state's propagation of punishment of government critics and rumour-mongers. The state suppresses government criticism to manipulate public opinion and make it favourable to those in power.⁷⁷ By highlighting the punishment of criticism of rank-and-file government officials and law enforcement officers, the authorities are saving face for the government and restoring its overall authority.

Similarly, the state's focus on rumours as one of the most intolerable online expressions reflects the state's increased concerns and anxieties over withering authority and public trust in the age of social media.⁷⁸ While the state primarily targeted high-profile social media influencers in the earliest anti-rumour campaigns,⁷⁹ our analysis shows that authorities began to target rumours more broadly and severely between 2014 and 2019. Although it is not immediately clear from our dataset how influential the targeted WeChat users were, official reports often stressed that it was the person who initiated the rumours, rather than the most influential ones who spread rumours, that faced the harshest punishment. Building on the analysis of Gallagher and Miller which focuses on the state targets of repression in 2012,⁸⁰ our results suggest that authorities have broadened their targets of repression from influential non-party thought leaders to almost anyone who dares challenge official narratives. According to Huang, rumours decrease citizens' trust in the government and support for the regime.⁸¹ It may therefore be the state's attempt to restore its credibility and legitimacy by discrediting and eliminating alternative sources of information and ensuring that citizens receive and rely on government-approved messages.

The larger number of reports featuring WeChat users' sharing of undesirable content in group chats and Moments rather than in one-to-one chats might be a result of users disproportionately posting sensitive content in different WeChat functionalities or the authorities purposefully targeting social media functionalities that reach larger audiences. In either case, that the Chinese authorities choose to propagate these cases of repression is notable and worth exploring. Existing work shows that the more ambiguous the boundary of online expression, the more likely average citizens self-censor to avoid potential trouble.⁸² There are two possible explanations for why the authorities choose to detail the reasons for each arrest even though they can maximize controls by blurring the bottom lines for online speech.

First, it serves as a deterrence tool. In Scott's words, 'each public punishment, each use of an honorific or a term of derogation – is a symbolic gesture of domination that serves to manifest and reinforce a hierarchical order'.⁸³ Public display of punishment is therefore a tool used by the powerful to convey deviant behaviours that the state would not tolerate and reflects what the state considers threatening behaviours among

average citizens. Second, it is a display of state power. In recent years, the CCP has been met with increasing scepticism about its regime stability and legitimacy.⁸⁴ Publicizing the state's action against dissenting views 'signals'⁸⁵ the state's strength and determination to control any speech or actions it deems undesirable. The Chinese state has been shown to target and make an example of high-profile social media users who dare challenge the authorities in public space,⁸⁶ leading to the conclusion that the state targets almost exclusively influential public opinion leaders. Our data shows that the state's controls over online information are more expansive than previously suggested. According to our data, the state has shown that it would not hesitate to go after any individuals who post undesirable content online and that it has omniscience over any online interactions, even those conducted in semi-public venues (e.g. instant messaging and WeChat Moments). Additionally, while not publicizing these behaviours may instil a sense of fear and uncertainty among average citizens, such action also runs the risk of making the state look weak or letting the masses think it is acceptable to challenge the authorities.⁸⁷ If individuals realize – through online close-knit channels such as WeChat's group chats and Moments – that they are not the only ones among their networks who harbour negative feelings towards the government or question the official narratives, then they might be emboldened to publicly challenge the regime. Detailing cases of repression shows the state's determination to control undesirable behaviours as well as its ability to do so.

Conclusion

China's online information control regime relies heavily on censorship, which is implemented by private companies. Because of such delegation, research on information controls drawing from results of social media censorship offers only a piece of the puzzle because they may not reflect accurately what the state wishes to control. Our analysis of when and why the state resorted to its repression apparatus to deter undesirable content furthers the understanding of China's information control regime from the perspective of the state.

Our study showed the state pushing the responsibility of content controls further down to the level of individuals. By detailing cases where individuals were punished for posting undesirable content on social media, the state's aim was to deter similar actions among the masses as well as demonstrate its strength.

We found that the state most frequently highlighted the consequences of posting content pertaining to government criticism and rumours. By making an example of individuals who challenge government actors, the authorities save face for the government and reiterate its authority. The goal of controlling rumours is likely to restore the state's credibility and its ability to guide public opinion by discrediting alternative narratives.

Our analysis also pointed to the need for further studies on China's information control strategies. For example, do certain types of content or users face harsher punishment than others? Are there any regional differences in the target and severity of state repression? Moreover, as propaganda and censorship are two sides of the same coin, future research can study censorship on WeChat during the same observation period of our study. By comparing what the state conceals and promotes around the same events, researchers may develop a better understanding of China's information control regime.

Notes

This study was made possible by funding from the Open Society Foundations. Research undertaken for this project was supervised by Ron Deibert. The authors thank the anonymous referees for their valuable comments.

1. Rebecca MacKinnon, China's censorship 2.0: How companies censor bloggers, *First Monday* 14(2), 2009, <https://journals.uic.edu/ojs/index.php/fm/article/view/2378/2089>, accessed 30 June 2020.
2. James C. Scott, *Domination and the Arts of Resistance: Hidden Transcripts*, New Haven, CT: Yale University Press, 1990.
3. Ervand Abrahamian, *Tortured Confessions: Prisons and Public Recantations in Modern Iran*, Berkeley: University of California Press, 1999; Mark Findlay, Show trials in China: After Tiananmen Square, *Journal of Law and Society* 16(3), 1989: 352.
4. Scripted and Staged: Behind the Scenes of China's Forced TV Confessions, Safeguard Defenders, April 2018, <https://safeguarddefenders.com/sites/default/files/wp-rsdl/uploads/2018/04/SCRIPTED-AND-STAGED-Behind-the-scenes-of-Chinas-forced-televised-confessions.pdf>, accessed 16 June 2020.
5. Qiuqing Tai, China's media censorship: A dynamic and diversified regime, *Journal of East Asian Studies* 14(2), 2014: 185–209.
6. Gary King, Jennifer Pan, and Margaret E. Roberts, How censorship in China allows government criticism but silences collective expression, *American Political Science Review* 107(2), 2013: 326–43.
7. Jedidiah R. Crandall et al., Chat program censorship and surveillance in China: Tracking TOM-Skype and Sina UC, *First Monday* 18(7), 2013, <https://firstmonday.org/ojs/index.php/fm/article/view/4628/3727>, accessed 30 June 2020.
8. Masashi Crete-Nishihata et al., Harmonized histories?: A year of fragmented censorship across Chinese live streaming applications, <https://netalert.me/assets/harmonized-histories/harmonized-histories.pdf>, accessed 16 June 2020.
9. Jeffrey Knockel, Lotus Ruan, and Masashi Crete-Nishihata, Measuring decentralization of Chinese keyword censorship via mobile games (paper presented at the 7th Usenix Workshop on Free and Open Communications on the Internet, Vancouver, 14 August 2017).
10. Blake Miller, The limits of commercialized censorship in China, 20 August 2019, <http://eprints.lse.ac.uk/id/eprint/101417>, accessed 30 June 2020.
11. Juha Antero Vuori and Lauri Paltmaa, The lexicon of fear: Chinese Internet control practice in Sina Weibo microblog censorship, *Surveillance & Society* 13(3/4), 2015: 400–21.
12. King, Pan, and Roberts, How censorship in China allows government criticism but silences collective expression.
13. Crandall et al., Chat program censorship and surveillance in China; Ashley Esarey and Xiao Qiang, Political expression in the Chinese blogosphere: Below the radar, *Asian Survey* 48(5), 2008: 752–72.
14. MacKinnon, China's censorship 2.0.
15. David Bandurski, The great hive of propaganda, China Media Project, 16 September 2017, <http://chinamediaproject.org/2017/09/16/the-great-hive-of-propaganda/>, accessed 16 June 2020.
16. Johan Lagerkvist, Principal-agent dilemma in China's social media sector?: The party-state and industry real-name registration waltz, *International Journal of Communication* 6, 2012: 2628–46.
17. Crandall et al., Chat program censorship and surveillance in China; Jeffrey Knockel et al., Every rose has its thorn: Censorship and surveillance on social video platforms in China (paper presented at the 5th Usenix Workshop on Free and Open Communications on the

- Internet, Washington, DC, 10 August 2015); Knockel, Ruan, and Crete-Nishihata, Measuring decentralization of Chinese keyword censorship via mobile games.
18. Miller, The limits of commercialized censorship in China.
 19. Lotus Ruan et al., One app, two systems: How WeChat uses one censorship policy in China and another internationally, Citizen Lab, 30 November 2016, <https://citizenlab.ca/2016/11/wechat-china-censorship-one-app-two-systems/>, accessed 16 June 2020.
 20. Perry Link, China: The anaconda in the chandelier, *New York Review of Books* 49(6), 2002, <https://www.nybooks.com/articles/2002/04/11/china-the-anaconda-in-the-chandelier/>, accessed 16 June 20.
 21. Ibid.
 22. Ruan et al., One app, two systems; Miller, The limits of commercialized censorship in China; Tai, China's media censorship; and Margaret E. Roberts, *Censored: Distraction and Diversion Inside China's Great Firewall*, Princeton, NJ: Princeton University Press, 2018.
 23. Kathleen Hartford, Dear mayor: Online communications with local governments in Hangzhou and Nanjing, *China Information* 19(2), 2005: 217–60; David Bandurski, Propaganda in the era of 'post-truth', China Media Project, 30 November 2018, <http://chinamediaproject.org/2018/11/30/propaganda-in-the-era-of-post-truth/>, accessed 16 June 2020; and Roberts, *Censored*.
 24. Jonathan Hassid, Safety valve or pressure cooker?: Blogs in Chinese political life, *Journal of Communication* 62(2), 2012: 212–30.
 25. Peter Lorentzen, China's strategic censorship, *American Journal of Political Science* 58(2), 2014: 402–14.
 26. Lynette H. Ong, 'Thugs-for-hire': Subcontracting of state coercion and state capacity in China, *Perspectives on Politics* 16(3), 2018: 680–95.
 27. Ibid.
 28. 新华社评薛蛮子嫖娼被拘: 伪善面具破碎斯文扫地 (*Xinhua* commented on Xue Manzi's detention on a charge of visiting prostitutes: Hypocrisy exposed and wallowing in degeneracy), 29 August 2013, <http://news.sina.com.cn/c/2013-08-29/005128073331.shtml>, accessed 30 June 2020.
 29. Jason Q. Ng, Politics, rumors, and ambiguity: Tracking censorship on WeChat's public accounts platform, Citizen Lab, 20 July 2015, <https://citizenlab.ca/2015/07/tracking-censorship-on-wechat-public-accounts-platform/>, accessed 16 June 2020.
 30. Jing Zeng, Chung-hong Chan, and King-wa Fu, How social media construct 'truth' around crisis events: Weibo's rumor management strategies after the 2015 Tianjin blasts, *Policy & Internet* 9(3), 2017: 297–320.
 31. Hu Yong 胡泳, 謠言作為一種社會抗議 (Rumour as social protest), *傳播與社會學刊* (Communication and society), no. 9, 2009: 67–94.
 32. James C. Scott, *Weapons of the Weak: Everyday Forms of Peasant Resistance*, New Haven, CT: Yale University Press, 1985.
 33. Haifeng Huang, A war of (mis)information: The political effects of rumors and rumor rebuttals in an authoritarian country, *British Journal of Political Science* 47(2), 2017: 283–311.
 34. Rogier Creemers, Cyber China: Upgrading propaganda, public opinion work and social management for the twenty-first century, *Journal of Contemporary China* 26(103), 2017: 85–100.
 35. Min Jiang, The business and politics of search engines: A comparative study of Baidu and Google's search results of Internet events in China, *New Media & Society* 16(2), 2014: 212–33.
 36. Chris Buckley, Crackdown on bloggers is mounted by China, *New York Times*, 10 September 2013, <https://www.nytimes.com/2013/09/11/world/asia/china-cracks-down-on-online-opinion-makers.html>, accessed 20 June 2020.

37. 关于办理刑事案件收集提取和审查判断电子数据若干问题的规定 (Provisions on several issues concerning the collection, obtaining, examination, and judgment of electronic data in the handling of criminal cases), 21 September 2016, https://www.spp.gov.cn/zdgz/201609/t20160921_167425.shtml, accessed 30 June 2020.
38. Public security administration punishment law of the People's Republic of China, 2 December 2016, <https://www.cecc.gov/resources/legal-provisions/public-security-administration-punishment-law-chinese-text>, accessed 16 June 2020.
39. Stanley Lubman, Arrested, detained: A guide to navigating China's police powers, *Wall Street Journal*, 12 August 2014, <https://blogs.wsj.com/chinarealtime/2014/08/12/arrested-detained-a-guide-to-navigating-chinas-police-powers/>, accessed 16 June 2020.
40. Ng, Politics, rumors, and ambiguity.
41. 具有舆论属性或社会动员能力的互联网信息服务安全评估规定 (Provisions on the security assessment of Internet information services having public opinion attributes or social mobilization capacity), <http://en.pkulaw.cn/display.aspx?cgid=d190ab08b1ae01dcbdfb&lib=law>, accessed 16 June 2020.
42. Mary Gallagher and Blake Miller, Who not what: The logic of China's information control strategy, 13 January 2019, <https://drive.google.com/file/d/1xjHWI0lh3abONxL0WoXj2swlcBuiRCXO/view>, accessed 30 June 2020.
43. Daniela Stockmann and Ting Luo, Which social media facilitate online public opinion in China?, *Problems of Post-Communism* 64(3–4), 2017: 189–202.
44. Scott, *Domination and the Arts of Resistance*.
45. Findlay, Show trials in China.
46. Abrahamian, *Tortured Confessions*.
47. Rachel E. Stern and Jonathan Hassid, Amplifying silence: Uncertainty and control parables in contemporary China, *Comparative Political Studies* 45(10), 2012: 1230–54.
48. Scott, *Domination and the Arts of Resistance*.
49. Corey Robin, *Fear: The History of a Political Idea*, New York: Oxford University Press, 2004.
50. Damien Ma and Neil Thomas, In Xi we trust: How propaganda might be working in the new era, MacroPolo, 12 September 2018, <https://macropolo.org/committee-analysis/in-xi-we-trust/>, accessed 16 June 2020.
51. Kecheng Fang and Maria Repnikova, Demystifying 'little pink': The creation and evolution of a gendered label for nationalistic activists in China, *New Media & Society* 20(6), 2018: 2162–85.
52. Roberts, *Censored*; Stockmann and Luo, Which social media facilitate online public opinion in China?
53. Anne-Marie Brady, *Marketing Dictatorship: Propaganda and Thought Work in Contemporary China*, Lanham, MD: Rowman & Littlefield, 2008.
54. Maria Repnikova and Kecheng Fang, Digital media experiments in China: 'Revolutionizing' persuasion under Xi Jinping, *China Quarterly* 239, 2019: 679–701.
55. Haifeng Huang, Propaganda as signaling, *Comparative Politics* 47(4), 2015: 419–37.
56. Scripted and Staged.
57. 全国首批50个省市网警执法账号将亮相网络空间 (First group of Internet police units at 50 locations including both provinces and municipalities will launch their own social media accounts), 中国新闻网 (China news), 31 May 2015, <http://www.chinanews.com/gn/2015/05-31/7312337.shtml>, accessed 16 June 2020.
58. Ibid.
59. Ibid.
60. Christopher Cairns, Fragmented authoritarianism?: Reforms to China's Internet censorship system under Xi Jinping, 12 September 2016, <http://www.chrismcairns.com/>

- uploads/3/0/2/2/30226899/cairns_--_fragmented_authoritarianism_--_reforms_to_chinas_internet_censorship_system_under_xi_jinping.pdf, accessed 10 August 2020.
61. Gallagher and Miller, Who not what.
 62. Stockmann and Luo, Which social media facilitate online public opinion in China?
 63. King, Pan, and Roberts, How censorship in China allows government criticism but silences collective expression.
 64. Crete-Nishihata et al., Harmonized histories?
 65. Zeng, Chan, and Fu, How social media construct ‘truth’ around crisis events.
 66. Ng, Politics, rumors, and ambiguity.
 67. Vuori and Paltemaa, The lexicon of fear.
 68. Lorentzen, China’s strategic censorship.
 69. Vuori and Paltemaa, The lexicon of fear; Miller, The limits of commercialized censorship in China; Ng, Politics, rumors, and ambiguity; and Esarey and Xiao, Political expression in the Chinese blogosphere.
 70. Link, China; MacKinnon, China’s censorship 2.0.
 71. Miller, The limits of commercialized censorship in China.
 72. Esarey and Xiao, Political expression in the Chinese blogosphere; Vuori and Paltemaa, The lexicon of fear; and Knockel et al., Every rose has its thorn.
 73. Ng, Politics, rumors, and ambiguity.
 74. 网民发‘习包子’被行拘引热议后 官方删除相关新闻 (Authorities deleted relevant news after discussions about arrest of a netizen who posted ‘Xi bun’ went viral), 24 August 2017, <https://bit.ly/2XWIdPc>, accessed 20 June 2020.
 75. King, Pan, and Roberts, How censorship in China allows government criticism but silences collective expression.
 76. Miller, The limits of commercialized censorship in China.
 77. Hartford, Dear mayor; Bandurski, Propaganda in the era of ‘post-truth’; Roberts, *Censored*; and Stockmann and Luo, Which social media facilitate online public opinion in China?
 78. Huang, Propaganda as signaling.
 79. Buckley, Crackdown on bloggers is mounted by China.
 80. Gallagher and Miller, Who not what.
 81. Huang, Propaganda as signaling.
 82. Link, China; Ng, Politics, rumors, and ambiguity.
 83. Scott, *Domination and the Arts of Resistance*.
 84. David Shambaugh, The coming Chinese crackup, *Wall Street Journal*, 6 March 2015, <https://www.wsj.com/articles/the-coming-chinese-crack-up-1425659198>, accessed 16 June 2020.
 85. Huang, Propaganda as signaling.
 86. Buckley, Crackdown on bloggers is mounted by China; Gallagher and Miller, Who not what.
 87. Scott, *Weapons of the Weak*; Hu, 謠言作為一種社會抗議.

References

- Abrahamian, Ervand (1999) *Tortured Confessions: Prisons and Public Recantations in Modern Iran*. Berkeley: University of California Press.
- 网民发‘习包子’被行拘引热议后 官方删除相关新闻 (Authorities deleted relevant news after discussions about arrest of a netizen who posted ‘Xi bun’ went viral) (2017) 24 August. <https://bit.ly/2XWIdPc>, accessed 20 June 2020.
- Bandurski, David (2017) The great hive of propaganda. China Media Project, 16 September. <http://chinamediaproject.org/2017/09/16/the-great-hive-of-propaganda/>, accessed 16 June 2020.

- Bandurski, David (2018) Propaganda in the era of 'post-truth'. China Media Project, 30 November. <http://chinamediaproject.org/2018/11/30/propaganda-in-the-era-of-post-truth/>, accessed 16 June 2020.
- Brady, Anne-Marie (2008) *Marketing Dictatorship: Propaganda and Thought Work in Contemporary China*. Lanham, MD: Rowman & Littlefield.
- Buckley, Chris (2013) Crackdown on bloggers is mounted by China. *New York Times*, 10 September. <https://www.nytimes.com/2013/09/11/world/asia/china-cracks-down-on-online-opinion-makers.html>, accessed 20 June 2020.
- Cairns, Christopher (2016) Fragmented authoritarianism?: Reforms to China's Internet censorship system under Xi Jinping. 12 September. http://www.chrismcairns.com/uploads/3/0/2/2/30226899/cairns_-_fragmented_authoritarianism_-_reforms_to_chinas_internet_censorship_system_under_xi_jinping.pdf, accessed 10 August 2020.
- Crandall, Jedidiah R., et al. (2013) Chat program censorship and surveillance in China: Tracking TOM-Skype and Sina UC. *First Monday* 18(7). <https://firstmonday.org/ojs/index.php/fm/article/view/4628/3727>, accessed 30 June 2020.
- Creemers, Rogier (2017) Cyber China: Upgrading propaganda, public opinion work and social management for the twenty-first century. *Journal of Contemporary China* 26(103): 85–100.
- Crete-Nishihata, Masashi, et al. (2017) Managing the message: What you can't say about the 19th National Communist Party Congress on WeChat. Citizen Lab, 6 November. <https://citizenlab.ca/2017/11/managing-message-censorship-19th-national-communist-party-congress-wechat/>, accessed 6 June 2020.
- Crete-Nishihata, Masashi, et al. (n.d.) Harmonized histories?: A year of fragmented censorship across Chinese live streaming applications. <https://netalert.me/assets/harmonized-histories/harmonized-histories.pdf>, accessed 16 June 2020.
- Esarey, Ashley and Xiao, Qiang (2008) Political expression in the Chinese blogosphere: Below the radar. *Asian Survey* 48(5): 752–72.
- Fang, Kecheng and Repnikova, Maria (2018) Demystifying 'little pink': The creation and evolution of a gendered label for nationalistic activists in China. *New Media & Society* 20(6): 2162–85.
- Findlay, Mark (1989) Show trials in China: After Tiananmen Square. *Journal of Law and Society* 16(3): 352–9.
- 全国首批50个省市网警执法账号将亮相网络空间 (First group of Internet police units at 50 locations including both provinces and municipalities will launch their own social media accounts) (2015) 中国新闻网 (China news), 31 May. <http://www.chinanews.com/gn/2015/05-31/7312337.shtml>, accessed 16 June 2020.
- Gallagher, Mary and Miller, Blake (2019) Who not what: The logic of China's information control strategy. 13 January. <https://drive.google.com/file/d/1xjHWI0Ih3abONxL0WoXj2swlcBuiRCXO/view>, accessed 30 June 2020.
- Hartford, Kathleen (2005) Dear mayor: Online communications with local governments in Hangzhou and Nanjing. *China Information* 19(2): 217–60.
- Hassid, Jonathan (2012) Safety valve or pressure cooker?: Blogs in Chinese political life. *Journal of Communication* 62(2): 212–30.
- Hu, Yong 胡泳 (2009) 謠言作為一種社會抗議 (Rumour as social protest). *傳播與社會學刊* (Communication and society), no. 9: 67–94.
- Huang, Haifeng (2015) Propaganda as signaling. *Comparative Politics* 47(4): 419–37.
- Huang, Haifeng (2017) A war of (mis)information: The political effects of rumors and rumor rebuttals in an authoritarian country. *British Journal of Political Science* 47(2): 283–311.
- Jiang, Min (2014) The business and politics of search engines: A comparative study of Baidu and Google's search results of Internet events in China. *New Media & Society* 16(2): 212–33.

- King, Gary, Pan, Jennifer, and Roberts, Margaret E. (2013) How censorship in China allows government criticism but silences collective expression. *American Political Science Review* 107(2): 326–43.
- Knockel, Jeffrey, et al. (2015) Every rose has its thorn: Censorship and surveillance on social video platforms in China. Paper presented at the 5th Usenix Workshop on Free and Open Communications on the Internet, Washington, DC, 10 August.
- Knockel, Jeffrey, Ruan, Lotus, and Crete-Nishihata, Masashi (2017) Measuring decentralization of Chinese keyword censorship via mobile games. Paper presented at the 7th Usenix Workshop on Free and Open Communications on the Internet, Vancouver, 14 August.
- Lagerkvist, Johan (2012) Principal-agent dilemma in China's social media sector?: The party-state and industry real-name registration waltz. *International Journal of Communication* 6: 2628–46.
- Link, Perry (2002) China: The anaconda in the chandelier. *New York Review of Books* 49(6). <https://www.nybooks.com/articles/2002/04/11/china-the-anaconda-in-the-chandelier/>, accessed 16 June 20.
- Lorentzen, Peter (2014) China's strategic censorship. *American Journal of Political Science* 58(2): 402–14.
- Lubman, Stanley (2014) Arrested, detained: A guide to navigating China's police powers. *Wall Street Journal*, 12 August. <https://blogs.wsj.com/chinarealtime/2014/08/12/arrested-detained-a-guide-to-navigating-chinas-police-powers/>, accessed 16 June 2020.
- Ma, Damien and Thomas, Neil (2018) In Xi we trust: How propaganda might be working in the new era. MacroPolo, 12 September. <https://macropolo.org/committee-analysis/in-xi-we-trust/>, accessed 16 June 2020.
- MacKinnon, Rebecca (2009) China's censorship 2.0: How companies censor bloggers. *First Monday* 14(2). <https://journals.uic.edu/ojs/index.php/fm/article/view/2378/2089>, accessed 30 June 2020.
- Miller, Blake (2019) The limits of commercialized censorship in China. 20 August. <http://eprints.lse.ac.uk/id/eprint/101417>, accessed 30 June 2020.
- Ng, Jason Q. (2015) Politics, rumors, and ambiguity: Tracking censorship on WeChat's public accounts platform. Citizen Lab, 20 July. <https://citizenlab.ca/2015/07/tracking-censorship-on-wechat-public-accounts-platform/>, accessed 16 June 2020.
- Ong, Lynette H. (2018) 'Thugs-for-hire': Subcontracting of state coercion and state capacity in China. *Perspectives on Politics* 16(3): 680–95.
- 政策法规 (Policies and laws) (n.d.) 中华人民共和国国家互联网信息办公室 (Cyberspace Administration of China). http://www.cac.gov.cn/zcfg/fl/A090901index_1.htm, accessed 30 June 2020.
- 关于办理刑事案件收集提取和审查判断电子数据若干问题的规定 (Provisions on several issues concerning the collection, obtaining, examination, and judgment of electronic data in the handling of criminal cases) (2016) 21 September. https://www.spp.gov.cn/zd gz/201609/t20160921_167425.shtml, accessed 30 June 2020.
- 具有舆论属性或社会动员能力的互联网信息服务安全评估规定 (Provisions on the security assessment of Internet information services having public opinion attributes or social mobilization capacity) (n.d.) <http://en.pkulaw.cn/display.aspx?cgid=d190ab08b1ae01dcbdfb&lib=law>, accessed 16 June 2020.
- Public security administration punishment law of the People's Republic of China (2016) 2 December. <https://www.cecc.gov/resources/legal-provisions/public-security-administration-punishment-law-chinese-text>, accessed 16 June 2020.
- Repnikova, Maria and Fang, Kecheng (2019) Digital media experiments in China: 'Revolutionizing' persuasion under Xi Jinping. *China Quarterly* 239: 679–701.

- Roberts, Margaret E. (2018) *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton, NJ: Princeton University Press.
- Robin, Corey (2004) *Fear: The History of a Political Idea*. New York: Oxford University Press.
- Ruan, Lotus, et al. (2016) One app, two systems: How WeChat uses one censorship policy in China and another internationally. Citizen Lab, 30 November. <https://citizenlab.ca/2016/11/wechat-china-censorship-one-app-two-systems/>, accessed 16 June 2020.
- Scott, James C. (1985) *Weapons of the Weak: Everyday Forms of Peasant Resistance*. New Haven, CT: Yale University Press.
- Scott, James C. (1990) *Domination and the Arts of Resistance: Hidden Transcripts*. New Haven, CT: Yale University Press.
- Scripted and Staged: Behind the Scenes of China's Forced TV Confessions (2018) Safeguard Defenders, April. <https://safeguarddefenders.com/sites/default/files/wp-rsdl/uploads/2018/04/SCRIPTED-AND-STAGED-Behind-the-scenes-of-Chinas-forced-televised-confessions.pdf>, accessed 16 June 2020.
- Shambaugh, David (2015) The coming Chinese crackup. *Wall Street Journal*, 6 March. <https://www.wsj.com/articles/the-coming-chinese-crack-up-1425659198>, accessed 16 June 2020.
- Stern, Rachel E. and Hassid, Jonathan (2012) Amplifying silence: Uncertainty and control parables in contemporary China. *Comparative Political Studies* 45(10): 1230–54.
- Stockmann, Daniela and Luo, Ting (2017) Which social media facilitate online public opinion in China? *Problems of Post-Communism* 64(3–4): 189–202.
- Tai, Qiuqing (2014) China's media censorship: A dynamic and diversified regime. *Journal of East Asian Studies* 14(2): 185–209.
- Vuori, Juha Antero and Paltemaa, Lauri (2015) The lexicon of fear: Chinese Internet control practice in Sina Weibo microblog censorship. *Surveillance & Society* 13(3/4): 400–21.
- 新华社评薛蛮子嫖娼被拘: 伪善面具破碎斯文扫地 (*Xinhua* commented on Xue Manzi's detention on a charge of visiting prostitutes: Hypocrisy exposed and wallowing in degeneracy) (2013) 29 August. <http://news.sina.com.cn/c/2013-08-29/005128073331.shtml>, accessed 30 June 2020.
- Zeng, Jing, Chan, Chung-hong, and Fu, King-wa (2017) How social media construct 'truth' around crisis events: Weibo's rumor management strategies after the 2015 Tianjin blasts. *Policy & Internet* 9(3): 297–320.