Analysing everyday online political talk in China: Theoretical and methodological reflections

_____

Scott Wright (University of Melbourne)
Todd Graham (University of Groningen)
Yu Sun (University of Groningen)
Wilfred Yang Wang (Queensland University of Technology)
Xiantian Luo (University of Melbourne)
Andrea Carson (University of Melbourne)

Abstract

This article explores the theoretical and methodological challenges of collecting and analysing everyday online political talk in China, and outlines our approach to defining and coding such talk. In so doing, the article is designed to encourage further research in this area, taking forward a new agenda for online deliberation (Wright, 2012a), and supporting this important area of research.

Keywords: online deliberation, China, social media, methodology

Introduction

Over the past two decades, a great deal has been written about the potential of the internet to facilitate political talk and deliberation, and to strengthen the public sphere, be it in China (Yang & Calhoun, 2007; Rauchfleisch & Schafer, 2015), Australia (e.g. Bruns et al., 2010), or beyond (Papacharissi, 2002; Wright 2007). Everyday political talk is the foundation stone of the public sphere (Habermas, 1984) and fundamental to civic life because "through everyday political talk, citizens construct their identities, achieve mutual understanding, produce public reason, form considered opinions…" (Kim & Kim, 2008, p. 51). Political talk is, quite simply, crucial to healthy citizenship (Dahlgren, 2006, p. 282) because it facilitates political knowledge, engagement and opinion change (Price & Cappella, 2002; Huckfeldt et al., 2004) and can lead people to take, or call for, political actions (Graham et al., 2015, 2016). Everyday political talk can do this because it encourages shared perspective building or complementary agency: intersubjective processes whereby people link their personal ideas, issues, and actions with one another, cultivating political agency, solidarity and community (McAfee, 2000, p. 134; Fearson, 1998). Such talk can be pre- or proto-political; a latent or

standby resource; 'potentially political', important to the 'microdynamics of democracy' (Dahlgren, 2006, p. 282; Ekman & Amna, 2012, pp. 287-8); and can provide a "gateway toward the stirrings of a broader social consciousness" (Howe, 2012), creating a sense of public empowerment and voice (Coleman, 2013, pp. 219-220), and facilitate broader civic involvement.

In considering everyday online political talk, China is a particularly interesting case because:

> Digital media in China reflect the many contradictions of the Chinese society: rapid diffusion but glaring digital divides, significant economic freedom but strict political control, new opportunities for civic engagement but with pervasive surveillance. The mix of politics with market and the unique Chinese culture have created a multifaceted Internet, sometimes reinforcing while other times restructuring political and social inequalities… (Chen and Reese, 2015, p. 1).

While political debate and action is heavily controlled and restricted in China (Bamman, O'Connor & Smith, 2012; King et al., 2013, 2014; Fu et al., 2013a, b), everyday talk turns political in a wide variety of everyday online contexts (or third spaces, see Wright, 2012b; Wright et al., 2016) and in various political 'shades' from the 'obvious' formal political topics to the more ambiguous and difficult to detect that are evocative of *the political* and personal turn (Wright, 2012a, b; Graham et al., 2016). Moreover, some Chinese 'Netizens' use coded language and metaphors to make political points and arguments, adding another layer of complexity to the debate (Rauchfleisch & Schafer, 2015). In combination, this makes it difficult for censors who want to identify and control political talk. Thus we argue that the different hues of political talk that emerge from within the everyday, 'non-political' Chinese online sphere(s) are an important avenue for political debate. However, this complexity also makes it difficult for researchers who want to analyse everyday political talk in China. It is perhaps unsurprising, therefore, that the analysis of everyday Chinese online political talk, particularly in 'third spaces', remains scant.

In this article, we outline our approach to theorising and operationalizing everyday political talk in China, and reflect on some of the challenges that we faced in doing so. The article is organised into three principal sections. First, we set out why everyday political talk matters so much in China, and our theoretical approach to defining and identifying everyday online political talk in China. Second, we outline our approach to analysing the nature of everyday online political talk in China. Finally, we discuss the different methodological challenges that we faced during our analysis in 2015, and how we met them.

The 'space' of everyday online political talk in China

In the Chinese context, the potential for the internet to facilitate a freer form of public political communication has been celebrated for its diversity (Rauchfleisch & Schafer, 2015; Song et al., 2016), with several examples of online political communication and activism leading to substantive changes in the law or, at least, to weak and sometimes scattered publics from which a civic identity can be built (Yang, 2009, 2013, p. 16; Qiu & Chan, 2011; Jiang, 2012; Sukosed & Fu, 2013). Jiang (2010) conceptualises the Chinese internet as a 'sphere composed of diverse yet connected spaces where the influence of the state varies, thus creating disparate conditions for public deliberation'. Jiang goes on to draw a distinction between four types of online space in China 'extending from the core to the peripheries of authoritarian rule: central propaganda spaces, government-regulated commercial spaces, emergent civic spaces, and international deliberative spaces'. For Jiang (2010), democracy is not necessarily a precursor for public deliberation because people may be able to circumvent authoritarian control. Even where they do not, there might be a form of permitted authoritarian deliberation (He & Warren, 2011).

Early research by Huang (1997) and Qiu (2000) was less positive about the impact of the internet on political talk in China, finding that a mixture of strict discussion rules and censorship in Chinese bulletin board systems (BBS) led to limited political talk on 'hard topic discussion' about democracy and other political topics, a finding that was broadly confirmed by Qiu (2000). For Qiu, online political talk is: "contingent upon the institutional barriers set by the Chinese authorities between the domestic and foreign cyberspaces, between the apolitical arenas and the sphere of Open Platform Communications. The totality of these constraints is virtual censorship…" (2000, p. 4). In a similar vein, Qinglian (2008) argues that the regime has successfully kept people ignorant and neutered the potential for political organising. However, research has shown that there are variations in how companies enforce regulations (MacKinnon, 2009), which is indicative of the complex socio-technical environment for online political talk in China.

At the heart of these debates is what we might call the space of political talk. Damm, for example (2007, p. 276), criticises the 'mono-causal interpretation' and the narrow focus on political websites which "fails to take into account the thousands of other websites, forums, and blogs left untouched. These are not focused specifically on political issues, but nevertheless discuss essential societal developments in China". As Yang (2009, pp. 1-2) notes, though, this complexity is often over-looked on the altar of two competing narratives:

… one of control and the other of entertainment" and this creates a "misconception that because of governmental internet control, Chinese internet users do nothing but play. The real struggles of the Chinese people are thus ignored, and the radical nature of Chinese internet culture is dismissed. Yet, not only is internet entertainment not apolitical, but political control itself is an arena of struggle […] The most unorthodox, imaginative, and subversive ideas can be found in Chinese cyberspace. Authority of all kinds is subject to doubt and ridicule. Ordinary people engage in a broad range of political action and find a new sense of self, community and empowerment.

Similarly Chen and Reese (2015, p. 2) observe that the "focus on the more visible examples of top-down regulation and control understates how networked technologies have helped create new forms of civic engagement from the bottom up" while Rosen (2010, pp. 512-513) argues that there is a fluidity to the "state-society equilibrium [that] is based on a compromise between an empowered public and an endangered leadership". While there are undoubtedly differences between core and periphery, as noted by Jiang, the central point of contention between the more positive and pessimistic analyses is the extent to which the Communist Party of China (CPC) has either chosen to *allow* "greater civic and political speech freedoms" and "relaxe[d] its grip over political discourse in exchange for its own legitimacy and survival" (Jiang, 2010, np) or is simply *unable* to control online activism as it responds and adapts to new restrictions (Yang, 2009, p. 44).

To address these debates, we argue that the space of political talk is key (Wright et al., 2016) and that it is important to analyse everyday political talk as it emerges in 'third spaces' which are formally non-political and may or may not be geographically-focused online communities, but where political talk can emerge (Wright, 2012a, b). To this end, we chose to focus our research on one potential Chinese third space: an online 'lifestyle' discussion forum (or BBS) with a significant help function. Interestingly, the forum chooses to pre-moderate every post. In other words, every message is vetted to decide whether or not to post it. In a commercially-run forum with tens of millions of posts, this is an exhaustive and expensive task. Thus, while arguably it is towards the periphery in Jiang's terms it is, at the same time, close. Thus it is questionable just how far the periphery is from the centre. While it may be an 'emergent civic space', it is also a 'government-regulated commercial space', and this may curtail its civic potential. We can assume that the moderators have actively allowed discernible political talk.

Analysing political talk in non-political online spaces raises a series of methodological challenges, and it is to this that we now turn.

'The political' beyond the political

Defining the political in political talk is notoriously tricky. In recent years, there has been a shift to encourage more inclusive definitions of politics in the western democratic context. This is because people are, it is thought, increasingly disengaged from formal politics, choosing to participate in more lifestyle-focused ways that are not captured by traditional definitions (Bennett, 1998; Hay, 2002, 2007; Ekman & Amna, 2012). As we know, formal political talk in China is heavily regulated. However, there may be more freedom for the politics of the everyday; of *the political*. Much of the literature on China focuses on debates around formal politics, contention or counter-hegemonic activity (e.g. Yang, 2009, 2013). There is a distinction to be made, however, between political talk and contention. Political contention will almost certainly require political talk, but political talk does not have to be contentious in character. For example, is someone engaging in political contention when talking about her or his personal experience of financial hardship bringing up a child? We would suggest not. However, following Mansbridge (1999), it is political talk to the extent that the person links the issue from the private sphere to the public context (Graham, 2008; Graham & Harju, 2011). But even if the talk remains private, it may still have an important political function (Dahlgren, 2006). Indeed, what might be considered pre/proto-political talk in a western democratic context arguably *is* political in China because it could be interpreted as questioning the principle of party supremacy and may be considered contentious, or even subversive, by the Chinese state—by power.

This arguably necessitates a shift in emphasis from 'collective' or 'public good' political talk and actions to private, personal/individualist—and perhaps lifestyle-oriented—topics, the value of which is contested (see Rosen, 2010; Bennett, 2003). If it is correct to argue that the 'unique characteristics' of online activism in China make it "more likely to be episodic and spontaneous, often without formal organization" in the vein of connective action (Yuan, 2015, p. 223; Bennett & Segerberg, 2013), this places greater emphasis on how actions emerge through everyday political talk in the general public sphere and how the fermenting and fomenting that occurs there percolates into the strong public sphere (Wright, 2015; Graham et al., 2015, 2016). In turn, this would suggest that scholars need to turn (or renew) their attention to "the restructuring of centre-periphery relations, politicizing the demands of various sectors of society, and the struggle over the definition of the realm of the political" and must pay particular attention to "new forms of mediated communication as a discursive field or space in which competing discourses struggle for visibility and legitimacy" (Yuan, 2015, p. 224).

Bringing this together, the barrier for what counts as political talk and action is arguably lower, or at least different, in China. It falls into grey areas—or "boundary spanning" as O'Brien and Li (2006, p. 50) describe it—that "crosses between the legitimate and illegitimate" (Yang, 2009, p. 3). While people may disagree with this analysis, and the boundaries have blurry edges, such pre/proto-political talk is, we argue, even more important in China, and can be considered a political action in and of itself. In making this analysis, how to interpret the party supremacy principle is important. At its most extreme, any discourse that questions the party could be considered as subversive or dissident. But, as has been shown, dissent is permitted within certain parameters—it depends on how it is framed and whether it could be perceived as mobilising. Some criticism can be seen as helping the CPC.[1]

Identifying everyday online political talk in China

Graham (2008) has noted that identifying political talk in non-political spaces is like looking for needles in a haystack, and this helps to explain the lack of research (Wright, 2012a) in this area. In previous research we have quantified the amount of political talk in third spaces by reading a random sample of all posts (Graham and Wright, 2014). This exploratory research design enabled us to understand the volume and nature of political talk in such spaces and informed our subsequent approach. While this understanding has value, it means that coders read large amounts of material that is not political. To overcome this problem, we subsequently used keywords to identify political talk, which can be categorised into four groupings: politicians, political parties, political institutions and general terms, such as 'democracy' and 'politics' (Graham et al. 2015, 2016). While each mention was checked to ensure it occurred in a political context, an initial sorting by keywords made the process of identifying political talk easier. We were also concerned that using broad definitions of politics would lead some people to argue that the results were about the definition of politics rather than a realistic account of political talk.

For the Chinese context, we were concerned that in a non-political context it would simply be too resource-intensive to manually read messages looking for political talk. Initially, we used a broad list of keywords, including both political institutions and governance, and a range of socio-political issues such as smog and food scandals.[3] This approach was deliberately broad because we did not want to make too many assumptions and because we also wanted to understand what topics were *not* being discussed or published. We also included the encoded terms for the same topics where we could identify them.

Analysing the coded public sphere can be 'a game of cat and mouse': if authorities discover a coded word, it can be added to the sensitive list. We used a mixture of personal knowledge and dictionaries of keywords (Ng, 2013).[4] For example, one of our keywords, 'CCTV' (China Central Television), is sensitive, and within China 'CCAV' is often used to avoid censorship. As with many coded words, it also carries a satirical meaning used to communicate discontent, here with CCTV and its role as the 'throat organ' of the party state.

Analysing and assessing everyday political talk in China

After identifying political talk, the next question is how to analyse such talk. A number of studies on the nature of online political talk in China focus on measuring the deliberativeness of political talk as a means of determining the extent to which online platforms are conducive to (particular) conditions of deliberation (e.g. Lewis, 2013; Thimm et al., 2014; Zhou et al., 2008). The work of Habermas—especially his theory of communicative rationality and discourse ethics—has been highly influential (1984, 1987, 1989). For example, Song et al. (2016) undertook a large-scale (machine-learning based content analysis and social network) analysis of political debate on Chinese food safety on Weibo (the Chinese version of Twitter). They conclude that

> much of the political talk on Weibo is emotional, expressing anger, fear or sadness […] As such, the Weibo mediated communication space cannot easily be defended as occasioning rational, critical discussion leading to consensus, nor to instrumental, utilitarian engagement leading to a concrete policy outcome (Song et al., 2016, p. 532).

The Song et al. study is interesting because it shines a light on the use of expressive communication. While this is framed negatively in terms of theories of the deliberative public sphere, the authors still see some value in this talk:

> Users' expressive repertoires on Weibo do reveal something important about civic engagement that is increasingly personalized and based in life practices. Within Weibo-mediated discussion space, there are a lot of people talking politics […] expressing emotion in online political talk enables a wider array of voices and perspectives to be heard on Weibo. […] This kind of life-style political engagement is hard to explain in terms of instrumental or rational communication. Rather, it is a clear manifestation of 'expressive rationality (Song et al., 2016, p. 532).

While Song et al. do not really explain what they mean by expressive rationality, and their textual analysis is an automated sentiment analysis (which is a rather limited way to assess expressiveness), their work suggests a need to move beyond rational-critical understandings of deliberation. This is an issue that we attempt to address here.

First, as has been argued elsewhere, analysing everyday talk in informal online communicative spaces requires a more inclusive definition of what constitutes the *talk* in political talk (e.g. Graham, 2008; Coleman & Moss, 2012; Wright, 2012a). Privileging formal notions of deliberation ignores the realities of everyday political talk; the ways in which people talk politics in ways that make sense to them. The everyday reality of making sense of politics is typically rooted in people's personal, subjective experiences, which are often expressed not through rational-critical debate but through other (often expressive) communicative forms such as storytelling, using humour and complaining. We might expect this to be particularly true in Chinese third spaces where humour is routinized in the coded public sphere. For these reasons, if we apply Habermas-inspired "idealized, and arguably impossible criteria by which to measure deliberation" in online places such as Weibo or lifestyle forums we might "preclude a positive outcome at the outset" (Wright, 2012a, p. 12). Thus, in order to provide a better understanding of how people talk politics in everyday online spaces, we need to move beyond elite normative frameworks by taking into account other communicative forms and the expressive nature of everyday talk.

Second, we wanted to allow room for exploring other communicative forms used by Chinese citizens, providing a more comprehensive account of the nature of online political talk in the Chinese context. To this end, we developed a content analysis coding scheme consisting of two analytical levels (building on Graham, 2008). Level one (divided into two parts) operationalizes Habermas' concept of the public sphere, and was created to assess the deliberative quality of political talk. First we will present communicative form and process. Building on the Habermas concept, rational-critical debate requires that participants provide reasoned claims which are critically reflected upon: "people's public use of reason" to support their claims in political debates is very crucial to the public sphere (Habermas, 1989, p. 27). Coherence and continuity is also considered important to deliberation: participants should stick to the topic of discussion until some form of agreement or understanding is achieved. This content analysis coding scheme for the first part of level one, deliberation, is operationalized into six code groups, as described in Table 1.

Table 1: Coding scheme Level 1A (part 1, Deliberation)

Code Title	Definition
Argument (reasoned claim)	A comment that provides a reasoned claim—presence of justification (formal presence of causal structures).
Evidence use	a) *Fact/Source*: An argument that supports its claim by providing a fact or source. b) *Example/anecdotal evidence:* An argument that supports its claim by providing a relevant example, which may include historical events, current events, comparisons and analogies between events, and hypothetical examples. c) *Personal experience*: An argument that supports its claim by providing personal experience.
Assertion (non-reasoned claim)	A comment that provides a non-reasoned claim: lack of justification.
Thematic consistency (coherence)	A comment that is on topic, in line with the political topic under discussion within the thread.
Argumentative depth (continuity)	A comment that is part of an exchange of claims, which includes: a) Counter: a comment that provides a reasoned claim in which an alternative claim is proposed in response to a competing claim or argument. b) Rebuttal: a comment that provides a reasoned claim in directly contradicting or challenging a competing claim or argument. c) Refute: a comment that provides a reasoned claim which directly defends an earlier claim or argument against a corresponding rebuttal. d) Affirmation: a comment that provides reasoned support in favour of another claim or argument.
Convergence	A comment that assents, concedes (partial assent), or agrees-to-disagree with/to another participant's claim or argument.

The second part of level 1, as shown in Table 2, focuses on the dispositional requirements for achieving mutual understanding. First, reciprocity requires that participants listen and respond to each other's questions, arguments, and opinions. Second, reflexivity is the internal or subjective process of reflecting another's argument or position against one's own. Third, an empathetic exchange occurs when a person tries to see themselves in another's shoes. Finally, sincerity implies that participants make all information relevant to the discussion (including their intentions, motives, desires and interests) known to other participants, and that all information provided is sincere and truthful

Table 2: Coding scheme Level 1B
(part 2, Dispositional requirements for achieving mutual understanding)

Reciprocity	All comments are coded for whether they are a reply to another post or are a stand-alone post: a) Stand-alone post: A comment that is not directed at any of the participants or other posts in the discussion. b) Reply: A comment that is a direct reply to another participant(s) or contents of a participant's post (typically done via the reply function).
Reflexive argument	A comment that provides: (a) a reasoned claim, an argument; (b) evidence to support that argument; (c) reasoned responsiveness to challenges by providing rebuttals and refutes; (d) and evidence in support of a challenge or defence against one.
Empathetic exchange	A comment that indicates the author has imagined his- or herself in another participants place/position. For example: "I understand where you are coming from".
Questionable sincerity	A comment that questions the sincerity or truthfulness of another participant's person, claim, argument, or statements in general.

The second level of the coding frame focuses on norms of debate (see Table 3). Discursive equality requires that participants respect each other as equals thereby prohibiting abusive and degrading communicative practices. Discursive freedom requires that participants are free to introduce and challenge assertions.

Table 3: Coding Scheme Level 2
(Norms of debate: Discursive equality and freedom)

Degrading	A comment that degrades—to lower in character, quality, esteem, or rank—another participant and/or participant's claim, argument or opinion in general.
Neglected	A comment containing an argument that is not attended to by other participants—lacking a reciprocal exchange.
Acknowledgement	A comment that acknowledges the presence, departure, or conversational actions of another participant, such as greeting, thanking, apologizing, and complementing.
Curbing	A comment that attempts to suppress, restrict, or prevent another participant's claim, argument, position, opinion, or statements in general.

As Table 4 reveals, level two of the coding scheme moves beyond formal notions of deliberation to identify other communicative forms and speech acts: attention, complaining, questioning, storytelling, and advice giving. First, the expression of attention is often used by Chinese internet users to convey their concern about a public issue. This can include comments like 'I will pay attention to this issue' or 'I will continue to pay attention to it'. In this way they are communicating their concern and promising to monitor an issue without calling for collective action or active intervention, such as protests, that would likely raise the ire of censors. Second, given the increasing social tensions in China, complaining and questioning are speech acts that contain much civic value. Through questioning, speakers can draw attention to the legitimacy of a certain policy or the authorities' way of dealing with problems, applying a pressure to act or, at least, to be accountable. Third, people engage in political talk by sharing personal experiences and stories (storytelling and chatter). Everyday political talk, especially in spaces dedicated to lifestyle issues, is often deeply connected to participants' personal lives. Such communicative practices open up spaces of personal and emotional relationships through which Chinese citizens can forge affective bonds that allow for deeper levels of understanding (Graham et al., 2015, 2016). Finally, we coded for advice giving, a form of civic involvement that potentially fosters a sense of belonging and community, thus strengthening the public sphere.

Table 4: Other communicative forms and speech acts

Attention	A comment that expresses concern about an issue. The symbol or word for attention is often used by people on Chinese social media and implies monitoring power from the citizens. For example: I will pay attention to this issue or continue to pay attention to it.
Complaining	A comment expressing a participant's dissatisfaction with an issue or certain state of affairs.
Questioning	A comment that poses questions concerning the issue or relevant policies. This includes comments that raise questions about the legitimacy of a policy or response (accountability).
Storytelling and Chatter	A comment where a participant tells a story (e.g. personal experience), gives an account of events (e.g. what they did that day) or simply provides some sort of personal information (e.g. likes, dislikes, interests).
Advice Giving/Helping	A comment that provides advice, recommendation or, more generally, helps another participant.

Having outlined both the theoretical background to, and method for, identifying and analysing everyday political talk in China, we next outline and address several challenges that we faced during the conduct of this research.

Methodological challenges

Our attempt in 2015 to analyse everyday political talk in China generated a number of challenges, and it is important to discuss these, and how we addressed them, as this may help others.

Issue 1: Collecting Weibo data

To analyse microblogs such as Weibo, the most common approach is to use an Application Programming Interface (API). Public APIs are created by websites to provide a 'window' onto some of their data in the 'back end', and to allow third party App developers an easy way to work with websites. They are particularly important for social media, where data is voluminous and dynamic. For researchers, APIs make data relatively easy to collect. However, there are many issues. First, many APIs are 'black boxes' and researchers do not know how representative the given data is. Thus, researchers need to think carefully about sample design and clearly communicate its limitations. Second, APIs typically do not provide access to historical data, often only going back a short period. Thus, researchers often collect live data, working from a point going forward. Third, at any time companies can change what data is made available through the API, and the general trend has been towards more restrictions – seemingly informed by an economic rationale that has similarities between Twitter and Weibo (Fuchs, 2015).[1]

Collecting Weibo data is more problematic. First, we were either refused access to Sina's APIs, or access was heavily restricted, to the point of being redundant. This meant that we could not extend an experimental Weibo data importer for NodeXL (a plug-in for Excel that can collect social media data and create social network maps). As we needed to use keyword searches to identify political talk, we chose to build a custom scraper to collect search returns directly from the website. Rather than asking for back end data through an API, scrapers work by visiting webpages and collecting data from the front end – from the visible website. Scraping is generally considered to be more difficult because, while most websites are written in a common language (e.g. HTML), the person(s) who construct or write the HTML all 'speak' in slightly different ways – thus websites typically have slightly different underlying code. This means that each website, and often different types of pages within a website, requires a unique script to be written to scrape the content. There are now tools that can scrape websites without the need to write code, such as Outwit Hub. Nevertheless, the process of visiting each page is time consuming and also meant that some of the back-end metadata available through the API could not be collected. Ultimately, though, we were successful in creating a useful sample.

Issue 2: Chinese BBS - scale and access

To collect data from the BBS, we again built a custom scraper. For this research we used a mixture of Outwit and custom-written scripts. However, collecting data from our Chinese forum proved difficult: numerous barriers were placed in our way.

Initially, we had wanted to analyse the impact of 'super-participants' on the nature (discursive equality) of political talk: 'super-posters' who had created more than 2000 posts (SP1s); set the agenda for debate (SP2s), and the moderators and facilitators (SP3s) – (see Graham and Wright 2014, 2015). In theory, the analysis covers all users and all threads since a forum was started, though often data is lost, deleted, or older threads are removed to ensure a forum does not slow down. Nevertheless, analysing super-participation proved very difficult. First, no overarching forum statistics were published (e.g. total numbers of posts created, users, threads), and this made it hard to know the scale of the task that we faced. This is important when considering sampling and research design. It was possible to collate some of this information by collecting all of the visible threads and the numbers of posts in each thread. However, as noted, this is unlikely to be all of the data because forums 'clean' or 'hide' older threads to maintain performance. Second, we found that the forum was often tortuously slow to load pages, and occasionally stopped returning data completely. This

meant we had to slow down data collection. Third, there was no list of members, something that is provided by most forums outside of China. To provide an over-arching analysis of super-participation, we needed aggregate statistics on the total numbers of threads started and posts created by each member. Each member does have an individual profile page from which, in theory, we could collect the data. In previous forums we have analysed (outside of China), the user-list has a common structure, going up sequentially with a unique URL for every account. In our Chinese BBS, profile pages were randomly assigned a URL within a large range. In an attempt to overcome this, we built another custom profile scraper and used a team of virtual machines in the cloud (that is, they were software-based implementations of a machine/computer on remote servers). We broke the URLs into a series of shorter lists and then visited each page to see if it was a valid account, collecting the username (for data verification) and total number of posts and threads created to see if they were super-posters (SP1s) or agenda-setters (SP2s).

Unfortunately, the combination of a slow website and the randomisation of user/member URLs left us insufficient time to collect the data. In total, we visited 3.1m pages, identifying 380,000 valid profile pages—identifying over 300 SP1s. As we did not have a complete list of users, we had to change our sampling design for this part of the study to a random sample of visible threads and then analyse the patterns of participation within this.

This speaks to another issue: the sheer scale of Chinese social media. Looking just at our incomplete list of SP1s, in the Chinese forum they had made a total of 4.03m posts, at an average of 5,036 posts each. Within the SP1s, 80 users had made over 10,000 posts, with the ten most frequent posters averaging 24,638 posts each. By comparing the number of posts created by SP1s with all the other users identified through the data collection, we found that SP1s had created 54% of all the posts, broadly in line with Graham and Wright (2014). But this is arguably distorted, because at the other extreme we found that 75,349 accounts -55% of all accounts—had never made a single post, while 17,454 users had made 1 post. Put simply, the more posts and users, the more resources and time are required to collect and work with the data.

Issue 3: Interface Design

The structure of website interfaces impacts the nature of the discussion online (Wright & Street 2007). Initially, we hoped to compare an Australian and a Chinese discussion forum with similar topics and a prima facie similar interface, alongside two ostensibly similar

micro-blogging platforms: Weibo and Twitter. However, there are a number of important differences that made this difficult/impossible.

Focusing on the micro-blogs first, arguably the structure of Weibo is better at facilitating deliberation than Twitter, and there is sometimes significant discussion on Weibo, with thousands of comments. Second, while Weibo has the same 140 character limit as Twitter, in Mandarin each character (or symbol) is a word and consequently Weibo allows a more extended comment and debate than Twitter (Rauchfleisch & Schäfer, 2015, p. 142). Weibo also allows the posting of videos, music, polls and long blogs. Weibo allows hashtags in a similar manner to Twitter, although there appears to be an approval process in place and this is generally more restricted. Another difference is that Weibo provides a detailed, hashtag-organised directory.

In regard to the Chinese discussion forum, because participants are told their comment is being reviewed complete and unsubtle pre-moderation encourages self-censorship. It seems likely that the lack of a public memberlist is a deliberate decision too. While this might be explained by a desire to protect user data, commercial forums (at least outside of China) provide such a list because it can help with community building. Another interpretation emphasises power and persuasion: a public list of users could be considered problematic because it makes it easier for people to identify social influencers and organise.

An important aspect in the popularity of Weibo is 'Big Vs' who are highly popular and influential verified users—often celebrities—and a key part of the company's marketing strategy. However, the potential political influence of 'Big Vs' – alongside the broader range of political talk occurring on Weibo (Sukosed & Fu, 2013; Yang, 2013) was an area of concern for the CPC. Big Vs were subjected to a crackdown. Several were arrested on a range of often unrelated charges (e.g. Xue Manzi), while new 'rumour laws' made it illegal to spread rumours, which are defined by the administration (Reuters, 2013). These two strategies proved quite effective at limiting politically sensitive talk. At the same time, many apparently state-sponsored pro-government Big Vs pushed out messages in support of the party.

Issue 4: The 50 Cent Party and 'internet pushing hands'

Businesses, lobby groups and governments all attempt to influence—and sometimes manipulate—online public opinion. For example, some bloggers secretly accept gifts or payment to positively review products. In our Chinese discussion forum, there is a whole

(public) section devoted to brands. The analysis of SP1s and SP2s indicated that a number were brand communication professionals.

There are also attempts to 'astroturf' political debate by creating content that supports a certain position and/or attempts to silence or distort critics. Research suggests that astroturfing does impact public opinion (Cho et al., 2011). Astroturfing would appear to be of a different order in China: members of the so-called '50 Cent Party' (*wu mao dang*), who are largely paid commenters, astroturf political talk. This means that the political talk we analyse may not be 'natural', but part of a strategic manipulation, and this is something that must be considered.

Issue 5: Analysing moderation and censorship

Moderation and censorship are crucial to the nature of online political debate (Wright, 2006). Within the model of super-participation, the third category (SP3s) is moderators and facilitators. When the typology was developed, we did not have authoritarian contexts in mind. Internationally most forums employ moderators or facilitators and normally one of their roles is to delete comments that do not meet community guidelines (e.g. foul language, trolling), and to more broadly encourage positive debate. Moderators and facilitators are, thus, an important influence on the nature of debate. China has a much more extensive, complex, and broadly government-directed system of political censorship and one of our research goals was to analyse how the Chinese moderators influenced debates. As most forums post-moderate, researchers can usually collect data as it is posted by users – before moderators have had chance to delete content – then compare this with what is left after moderation (King et al., 2013, 2014; Wright, 2006). However, our Chinese forum uses total pre-moderation, and thus we had to change our approach. In particular, the existence (or not) of published political talk (and coded political talk to avoid the censors—Esarey and Xiao, 2008), as well as the topics of that talk, becomes an important research question and one that we attempted to address. It is not so much about the nature of the talk, but how it is regulated (King et al., 2013, 2014). An alternative approach is to create accounts and post messages with different political topics to test what is deleted or blocked (e.g. MacKinnon, 2009). We decided not to use this method because of ethical concerns.

Conclusion

In this article we have set out our theoretical and methodological approach to analysing everyday political in non-political online Chinese 'third spaces', as well as the methodological challenges that we faced. In so doing, the article hopes to encourage and facilitate further research into everyday online political talk in China. We argue both for a more inclusive definition of politics to be adopted—one that includes everyday political issues as well as 'coded' political terms used to avoid censorship—and for the study of political talk in formally non-political online third spaces. In this context, we argue that researchers cannot just use Habermas-inspired models of rational critical communication, but also need to value and analyse expressive forms of communication, such as humour. We then set out our coding frame for analysing everyday political talk in China. Finally, we turn to the methodological issues that we faced when analysing Chinese social and digital media, and how we responded to them. In particular, there were issues with getting access to data, its mediated quality and the sheer scale of the data.

References

Bamman, D., O'Connor, B., & Smith, N. (2012). Censorship and deletion practices in Chinese social media. *First Monday, 17*(3). Retrieved from http://firstmonday.org/article/view/3943/3169

Bennett, W.l. (1998). The uncivic culture: Communication, identity, and the rise of lifestyle politics. Ithiel de Sola Pool Lecture, American Political Science Association, published in *P.S.: Political Science and Politics, 31*, 41-61.

Bennett, W.L. (2003). Communicating global activism: Strengths and vulnerabilities of networked politics. *Information, Communication & Society, 6*, 143-168.

Bennett, W.L., & Segerberg, A. (2013). *The logic of connective action*. New York, NY: Cambridge University Press.

Bruns, A., Burgess, J., Highfield, T., Kirchhoff, L., & Nicolai, T. (2010). Mapping the Australian networked public sphere. *Social Science Computer Review, 29*(3), 277-287.

Chen, W., & Reese, S.D. (Eds) (2015). *Networked China*. New York: Routledge.

Cho, C., Martens, M., Kim, H., & Rodrigue, M. (2011). Astroturfing Global Warming: It Isn't Always Greener on the Other Side of the Fence. *Journal of Business Ethics*, *104*(4), 571-87.

Coleman, S., & Moss, G. (2012). Under construction: The field of online deliberation research. *Journal of Information Technology & Politics, 9*(1), 1-15.

Coleman, S. (2013). *How Voters Feel*. Cambridge: Cambridge University Press.

Dahlgren, P. (2006). Doing citizenship: The cultural origins of civic agency in the public sphere. *European Journal of Cultural Studies*, *9*(3), 267-286.

Damm, J. (2007). The internet and the fragmentation of Chinese society", *Critical Asian Studies*, *39*(2), 273-294.

Deng, R. (2015). *Watchdog on extended party leashes? The paper mobile news app*. Public lecture: University of Melbourne.

Ekman J., & Amna E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, *22*(3), 283-300.

Esarey, A., & Xiao, Q. (2008). Political expression in the Chinese blogosphere: Below the radar. *Asian Survey*, *48*(5), 752-772.

Fearson, J.D. (1998). Deliberation as discussion. In J. Elster (Ed.), *Deliberative democracy* (pp. 44-68). Cambridge: Cambridge University Press

Fu, K.-w., Chan, C.-h., & Chau, M. (2013a). Assessing censorship on microblogs in China: Discriminatory keyword analysis and the real-name registration policy. *Internet Computing, IEEE, 17*(3), 42-50.

Fu, K.-w., & Chau, M. (2013b). Reality check for the Chinese microblog space: A random sampling approach. *PLoS ONE*, *8*(3). doi:10.1371/journal.pone.0058356

Fuchs, C. (2016). Baidu, Weibo and Renren: The global political economy of social media in China. *Asian Journal of Communication*, *26*(1), 14-41.

Graham, T. (2008). Needles in a haystack: A new approach for identifying and assessing political talk in nonpolitical discussion forums. *Javnost-The Public*, *15*(2), 17-36.

Graham, T., & Wright, S. (2014). Discursive equality and everyday political talk: The impact of super-participants. *Journal of Computer-Mediated Communication*, *19*(3), 625-642.

Graham T., & Harju, A. (2011). Reality TV as a trigger of everyday political talk in the net-based public sphere. *European Journal of Communication*, *26*(1), 18-32.

Graham, T., Jackson, D., &Wright, S. (2016). 'We need to get together and make ourselves heard': everyday online spaces as incubators of political action, *Information, Communication & Society, 19*(10)*,* doi:10.1080/1369118X.2015.1094113

Graham, T., Jackson, D., & Wright, S. (2015). From everyday conversation to political action: Talking austerity in online 'third spaces'. *European Journal of Communication, 30*(6), 648-665.

Graham, T., &Wright, S. (2014). 'Super-participation' in third spaces: volume and impact on political argument. In: R. Gibson, S. Ward & M. Cantijoch (Eds). *Analysing Social Media Data and Web Networks* (pp. 197-215). Basingstoke: Palgrave.

Habermas, J. (1984). *The theory of communicative action. (Vol. One), Reason and the rationalization of society*. Boston: Beacon Press.

Habermas, J. (1987). *The theory of communicative action. (Vol. Two), Life world and system: A critique of functionalist reason*. Boston: Beacon Press.

Habermas, J. (1989). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society* (trans. T. Burger). Cambridge: Polity Press.

Hay, C. (2002). *Political analysis: A critical introduction*, Basingstoke: Palgrave.

Hay, C. (2007). *Why we hate politics*. Cambridge: Polity.

He, B., & Warren, M.E. (2011). Authoritarian deliberation: The deliberative turn in Chinese political development. *Perspectives on Politics*, *9*(2), 269-289.

Howe, J.P. (2012). Slacktivism: A gateway, and only a gateway, to truly changing lives, *Huffington Post*. Retrieved from http://www.huffingtonpost.com/john-p-howe-iii-md/slacktivism-a-gateway-and_b_1422388.html

Huang, E.S. (1997). Flying freely but in a cage. Paper presented in the 1997 convention of *Association for Education in Journalism and Mass Communication*. Chicago, Illinois.

Huckfeldt, R., Mendez, J.M., and Osborn, T.L. (2004). Disagreement, ambivalence, and engagement: The political consequences of heterogeneous networks. *Political Psychology*, *25*(1), 65-95.

Jiang, M. (2010), Spaces of authoritarian deliberation: Online public deliberation in China. In E. Leib & B. He (Eds) *The search for deliberative democracy in China* (pp. 261-287). New York, NY: Palgrave.

Jiang, Y. (2012). *Cyber-Nationalism in China.* Adelaide: University of Adelaide Press.

Kim, J., & Kim, E.J. (2008). Theorizing dialogic deliberation: Everyday political talk as communicative action and dialogue. *Communication Theory*, *18*(1), 51-70.

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, *107*(2), 326-343.

King, G., Pan, J., & Roberts, M.E. (2014). Reverse-engineering censorship in China: Randomized experimentation and participant observation. *Science*, *345*(6199), 1–10.

Lewis, O.A. (2013). Net inclusion: New media's impact on deliberative politics in China. *Journal of Contemporary Asia*, *43*(4), 678-708.

MacKinnon, R. (2009). China's censorship 2.0: How companies censor bloggers. *First Monday*, *14:2*. Retrieved from http://firstmonday.org/article/view/2378/2089

Mansbridge, J. (1999). Everyday talk in the deliberative system. In S. Macedo (Ed.) *Deliberative Politics* (pp. 211-239). Oxford: Oxford University Press.

McAfee, N. (2000). *Habermas, Kristeva, and citizenship*. Ithaca, NY: Cornell University Press.

Ng, J. (2013). *Blocked on Weibo.* New York, NY: The New Press.

O'Brien, K., & Li, L. (2006). *Rightful resistance in rural China*. Cambridge: Cambridge University Press.

Papacharissi, Z. (2002). The virtual sphere: The internet as a public sphere. *New Media & Society*, *4*(1), 9-27.

Price, V., & Cappella, J. (2002). Online deliberation and its influence: The electronic dialogue project in Campaign 2000. *IT & Society*, *1*(1), 303-329.

Qinglian, H. (2008). *The fog of censorship: Media control in China*. New York, NY: HRIC.

Qiu, L., & Chan, T. (2011). *Studies of new media events*. Beijing: People's University Press.

Qiu, J.L. (2000). Virtual censorship in China: Keeping the gate between the cyberspaces. *International Journal of Communications Law and Policy*, *4*, 1-25.

Rauchfleisch, A., & Schafer, M.S. (2015). Multiple public spheres of Weibo: a typology of forms and potentials of online public spheres in China. *Information, Communication & Society*, *18*(2), 139-155.

Reuters (2013). China threatens tough punishment for online rumour spreading. *Reuters*. Retrieved, from http://www.reuters.com/article/us-china-internet-idUSBRE9880CQ20130909.

Rosen, S. (2010). Is the internet a positive force in the development of civil society, a public sphere, and democratization in China? *International Journal of Communication*, *4*, 509-516.

Simons, M., Nolan, D., & Wright, S. (2016, May online first). 'We are not North Korea': propaganda and professionalism in the People's Republic of China. *Media, Culture & Society.* doi: 10.1177/0163443716643154

Song, Y., Dai, X-Y., & Wang, J. (2016). 'Not all emotions are created equal': Expressive behavior of the networked public on China's social media site. *Computers in Human Behavior*, *60*, 525-533.

Sukosed, M., & Fu, K.-W. (2013). How Chinese netizens discuss environmental conflicts? , In *Framing and Functions on Sina Weibo* (pp. 1-19). Department of Media and Communication, City University of Hong Kong.

Thimm, C., Dang-Anh, M., & Einspänner, J. (2014). Mediatized politics-structures and strategies of discursive participation and online deliberation on Twitter. In A. Hepp & F. Krotz (Eds) *Mediatized Worlds*, (pp. 253-270), Basingstoke: Palgrave.

Twitter (2015). "The Search API". Retrieved from https://dev.twitter.com/rest/public/search

Wright, S. (2006). Government-run online discussion fora: Moderation, censorship and the shadow of control. *British Journal of Politics and International Relations*, *8*(4), 550-568.

Wright, S. (2007). A virtual European public sphere? The Futurum discussion forum. *Journal of European Public Policy, 14*(8), 1167-1185.

Wright, S. (2015). Populism and Downing Street e-petitions: Connective action, hybridity and the changing nature of organizing. *Political Communication*, *32*(3), 414-433.

Wright, S. (2012a). Politics as usual? Revolution, normalization and a new agenda for online deliberation. *New Media & Society*, *14*(2), 244-261.

Wright, S. (2012b). From 'third place' to 'third space': Everyday political talk in non-political online spaces. *Javnost: The Public*, *19*(3), 5-20.

Wright, S., & Street, J. (2007). Democracy, Deliberation and Design: The Case of Online Discussion Forums. *New Media & Society*, *9*(5), 849-869.

Wright, S., Graham, T., & Jackson, D. (2016). Third space, social media, and everyday political talk. In Bruns, A., Skogerbø, E., Christensen, C. Larsson, A-O., Enli, G. (Eds) *Routledge Companion to Social Media and Politics* (pp. 74-88), London: Routledge.

Yang, G., & Calhoun, C. J. (2007). Media, civil society, and the rise of a green public sphere in China. *China Information*, *21*(2), 211-236.

Yang, G. (2009). *The power of the internet in China: Citizen activism online*. New York, NY: Columbia University Press.

Yang, G. (2013). Contesting food safety in the Chinese media: between hegemony and counter-hegemony. *The China Quarterly*, *214*, 337-355.

Yuan, E.J. (2015). The new political of mediated activism in China: A critical review. In: W. Chen, & S.D. Reece (Eds) *Networked China* (pp. 215-231). New York, NY: Routledge.

Zhou, X., Chan, Y-Y., & Peng, Z-M. (2008). Deliberativeness of online political discussion: A content analysis of the Guangzhou Daily website. *Journalism Studies*, *9*(5), 759-770.

Endnotes

[1] A journalist from *The Paper*—a state-sponsored news website that has published stories on corruption—neatly sums up this balance: we have "never crossed the bottom line, neither has [The Paper] lost its background color–we have always been 'red'. The message we want to send to the leadership is The Paper is one of the family…" (Deng, 2015). The problem—be it for citizens or journalists—is that these boundaries are (deliberately) left vague and constantly moving (Simons et al., 2016, online first).

[2] For example, Twitter previously had a 1500 tweet limit, but increased this to a maximum of 18000 in March 2013 when changing from API 1.0 to API 1.1 but also effectively neutered access to the follows edge. For businesses, there is a balancing act: on the one hand, providing access to data through APIs can increase use, Apps, and thus positively impact profit. But the very data being 'given out' is also commercially valuable and selling access to the data (e.g. Twitter's Firehose) is an important part of the business model. Twitter's Search API is searchable across several criteria including hashtag and keyword, but only returns a limited number of tweets and focuses on 'relevance and not completeness' (Twitter, 2015). Furthermore, Twitter limits how often users can request data from the API (currently 180 times every 15 minutes), and if this is exceeded the user will be rate-limited—effectively paused—and the remaining data will be missed.

[3] The general keywords were clustered into topic areas with political institutions and governance (e.g. media institutions/ party media/CCTV or CCAV- 媒体/党媒/ CCTV or CCAV) alongside general topics such as parenting and childcare (which captured issues such as milk powder/food safety/children's health -奶粉/食品安全/儿童公共卫生安全) and left-behind rural children (留守儿童); marriage and family (including topics such as the marriage law/divorce law (婚姻法/离婚法) and having more than one child (二胎/三胎); public health (e.g. smoking/anti-smoking/ban on smoking 吸烟/禁烟/控烟); and the environment (e.g. smog and climate change 雾霾/气候变化).

[4] There are many existing dictionaries of words that can be used, such as the Stanford NLP Chinese database: http://nlp.stanford.edu/software/CRF-NER.shtml.

Authors

Scott Wright is a Senior Lecturer in Political Communication at the University of Melbourne. His research interests include everyday political talk, online deliberation, online civic engagement, democratic innovation, e-democracy and digital journalism.

Todd Graham is an Assistant Professor in Journalism and Political Communication at the Groningen Centre for Media and Journalism Studies, University of Groningen. His main research interests are the use of new media in representative democracies, the intersections between popular culture and formal politics, online election campaigns, online deliberation and political talk, and online civic engagement.

Yu Sun is a PhD candidate at the Center for Media and Journalism Studies at the University of Groningen, the Netherlands. Her research interests involve: public deliberation, civic engagement and the use of social media in the context of China.

Wilfred Yang Wang is a Research Fellow at the Digital Media Research Centre at Queensland University of Technology, Australia. His current research focuses on digital Asia and internet geography

Xiantian Luo is a senior software engineer, having graduated from University of Melbourne with a Masters in IT. He has extensive experience and has worked for multi-national tech companies. His research interests include internet crawling, information retrieval, big data, cloud computing and mobile application development.

Andrea Carson is a lecturer in Media and Politics at the University of Melbourne. She is also an honorary fellow at the University's Centre for Advancing Journalism. Her research interests include the relationship between news media and democracy, investigative and data journalism, digital media and political communication.