



Why do replies appear? A multi-level event history analysis of online policy discussions

new media & society

1–21

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1461444819880520

journals.sagepub.com/home/nms



Weiyu Zhang 

Yipeng Xi

National University of Singapore, Singapore

Anfan Chen

Department of Science Communication and Science Education, University of Science and Technology of China, China

Abstract

Participation takes time in both synchronous and asynchronous communication. To encourage participation, most scholars suggested strategies to change participants' motivation to adjust the way how they distribute their time resources. We, instead, view time as an evolving environment with specific temporal norms. This study employed a multi-level event history model to explore what factors affect response behaviors in an online policy discussion forum. By proposing four time-constant individual post factors and four time-varying environment factors, and analyzing 1798 posts in the forum, the study found that both individual post and environment factors are significant predictors of online interactions in different time phases. Some interaction effects between individual post and environment factors were also found. Rather than merely focusing on participants' resource to take part in online discussions, we suggest that future studies should examine how contextual constraints that change over time can influence response behaviors.

Keywords

Multi-level event history analysis, online discussion, reply, response, time-constant variables, time-varying variables

Corresponding author:

Weiyu Zhang, Department of Communications and New Media, National University of Singapore, Room 03-24, 11 Computing Drive, Singapore 117416.

Email: weiyu.zhang@nus.edu.sg

As computer-mediated communication (CMC) is widely applied in supporting human interactions, the asynchronous threaded forum becomes a typical application of CMC to encourage citizens to post messages and participate in interpersonal discussions at any time and from anywhere (Berge, 1997; Jeong, 2004). Such detachment of time and the human body is the initial aim that motivates the invention of asynchronous discussion platforms, which offer participants more time reading, reflecting, composing, and posting responses as well as justifying their rationales (Friess and Eilders, 2015). Although asynchronous discussions are conducive to quality interactions, they can also have potential drawbacks because participants must wait for replies over a period of time, which can restrict the momentum and flow of discussions (Hewitt, 2005; Jeong, 2004).

Participation and interactivity are key dimensions of online discussions. Discussion should be a social process of giving and taking, a mutual exchange of listening and responding (Barber, 1984: 175). Arguments should not only be articulated but also listened and replied to (Friess and Eilders, 2015; Zhang et al., 2013). Hewitt and Teplov's (1999) study on this theme showed that discussion threads could be negatively affected by long response time, but other studies (Hewitt, 2005; Jeong, 2004) suggested that the content or function of messages should not be neglected in determining how likely it is that a thread will maintain its vitality. However, few studies (Jeong, 2004; Jeong and Frazier, 2008) to date have examined the combined effect of individual posts and environment on the growth of discussion threads, and even fewer studies (e.g. Choi et al., 2018) have examined the explanatory factors of response behaviors from a temporal perspective. Therefore, this article theorizes and redefines the notion of temporality in online discussions and contributes further to strategies for stimulating interactivity in online discussions.

This article regards time as a changing scene or an evolving environment and starts with a brief review of the dynamic relationship between time and participation in online discussions. Situated in the context of online policy discussions, we then introduce eight hypotheses and one research question based on the existing theoretical arguments and empirical evidence. In addition to individual post factors that are constant over time, environment variables such as the crowdedness, quality, sentiment, and diversity of the forum discussion that change over time have to be included in our analysis. Using the method of multi-level event history analysis, we try to explain how both time-constant and time-varying factors prompt participants to reply to a post. We find that the time-varying environment factors function differently from the time-constant individual post factors. Moreover, the environment factors can moderate the relationship between the post factors and whether a post gets replied to or not.

Literature review

Why time matters in social life

Time is foundational to human life but is rarely studied by social scientists. Often, the assumption about time is that it is universal (e.g. 1 minute is 1 minute to everybody): always running short (e.g. one can only get older) and independent of human influence (e.g. nobody can change 1 day into 25 hours). Based on this assumption, time is

a precious resource that one has to manage carefully. This resource perspective emphasizes the use of time based on one's preference and availability, which have been enhanced by the advanced communication and transportation technology (Howlett and Goetz, 2014). Because of the ever-increasing autonomy of individuals, time understood as "time owned by oneself" has been emphasized more. Complaints like "bad timing," "busy doing something," and "undue haste" are examples to show how time as an individual resource can pose challenges to human interactions (Howlett and Goetz, 2014). This article takes a different perspective elaborated by Lefebvre (2004), which suggests that social life is organized around temporal norms and the temporal norms are culturally and interpersonally constructed, more than being controlled by individuals. Our article aims to explain how such temporal norms are developed in an online policy discussion forum.

Lefebvre is not alone in seeing time as being collectively constructed. The cultural influence on time has been discussed in Hall's (1983) work. For instance, when someone perceives time as a commodity for his or her exclusive use, it may indicate that she or he comes from an individualist society. Hall (1983) elucidated that time use patterns represent the "primary-level culture" (p. 6) that one lives in, and the temporal norms denote a set of underlying, hidden, unspoken, and implicit rules of behavior and thought. Walther's hyperpersonal communication theory stipulates that with enough time, CMC is able to provide a relationship that is as good as those established in a face-to-face setting. This theory suggests that time is one of few interpersonal cues that may convey rich meanings in the "cues-filtered-out" CMC environment. For example, individuals who let others wait for a reply for a long time are usually perceived as taking superior roles (Walther and Tidwell, 1995). In this sense, temporal cues could reveal one's position in the social hierarchy.

Lefebvre (2004) further argued that temporal norms are not fixed (e.g. one culture must have one norm all the time), but rather they are dynamic and heterogeneous. Such dynamic temporal norms are constructed through constant interactions among multiple rhythms. Lefebvre (2004: 6) proposed three types of interactions among rhythms: polyrhythmia, eurhythmia, and arrhythmia. Polyrhythmia refers to a variety of distinctive rhythms coexisting in a particular context, without much interaction. Eurhythmia and arrhythmia, then, denote the harmonious and conflicting conditions when various rhythms interact. Eurhythmia is an ideal circumstance, but it is a rarer state compared with arrhythmia. In the case of an online discussion forum, we can expect multiple rhythms to emerge at the beginning (polyrhythmia), and as the discussion continues, a development of both conflicts (arrhythmia) and integration (eurhythmia) may follow. We thus aim to examine the interactions among various rhythms, seeing time as a changing scene or an evolving environment that develops its own temporal norms that a participant needs to follow.

Time in online discussions

Online discussions take many forms (e.g. social media, instant messaging, etc.) and in this article, we define online discussions as an exchange of ideas among multiple individuals, in both synchronous and asynchronous manners. Specifically, we designed an electronic platform to stimulate the sharing of ideas, encourage engagement of

viewpoints, and elicit collective decision-making, within a pre-set time limit (8 days). The interactions among different time use patterns in such an online discussion forum could be observed in the following four aspects, namely, latency, sequence, density, and frequency.

Latency. Latency refers to the lag time between posts in a discussion forum (Dringus and Ellis, 2010). While asynchronous tools are flexible in terms of responding to others, too much lag time between messages can be perceived negatively by participants as they generally want to be heard by others or be acknowledged for their views (Dringus and Ellis, 2010). The long response time may also hinder the growth of discussion threads. Different types of message content were found to lead to different amounts of waiting time for replies (Hewitt and Teplovs, 1999). While posts that expand on an idea provided by others require the least amount of waiting time, those establishing arguments and finding flaws in others' responses need a longer lag time to respond to (Jeong, 2004). A dilemma seem to exist: short lag time is good for the growth of discussion threads, but high-quality replies need more time to respond.

Sequence. Time may also lead to a sequencing problem because the arrow of time goes forward at least in human psychology. As asynchronous communication restricts nonverbal components, people may find it hard to start or terminate a conversation (Chen et al., 2017). As time goes by, the earlier posts are pushed down in the forum threads. Although the storage capacity of computers permits new participants to review previous posts (Majchrzak and Malhotra, 2016), it is always troublesome for newcomers to go back to all previous content, and more importantly, previous posts may be displayed in an order that is chronological rather than logical to the flow of discussions. Therefore, new participants may start talking about an issue that has already been discussed and got ignored. The difficulty in tracking the logical order of discussions may result in the loss of vitality of online discussions.

Density. Density, or pace, refers to the number of events that occur during a specific period of time (Dringus and Ellis, 2010; Hesse et al., 1988). A higher-density discussion involves the exchange of more cues, words, and opinions than a lower-density discussion within a given unit of time. It also requires participants to prepare well for the discussion ahead of time and focus on the given task when the discussion starts, which increases the efficiency of collaborative working (Durrington and Yu, 2004). Contrary to the above findings, Jeong and Frazier (2008) argued that as those arguments and critiques require more time to digest and reflect, a high-density discussion may occupy the attention that should have been allocated to the exchange of high-quality views. Moreover, unless being encouraged, most participants prefer to maintain superficial and limited connections with their peers, leading to a low-level of density during the discussion period (Macfadyen and Dawson, 2010).

Frequency. Frequency refers to the number of posts/active users at a given time of the day, or day of the week. It differs from density because frequency focuses on recurring patterns. While density hypothetically can range widely, frequency is supposed to

identify if there are repeated events, depending on, for example, which day of a week it is. Most studies found that posts made early in a week may receive a larger number of responses than those on other days (Jeong and Frazier, 2008; Medaglia and Yang, 2017; Tatsumi and Nakazawa, 2017). In addition, the number of posts/active users usually reaches its peak in the afternoons or early evenings (Gibbs et al., 2008). The above findings indicate that people have a time preference in terms of participating in online discussions, though whether the preferences can lead to replies also depends on the density of postings and the depth of content in the forum (Jeong and Frazier, 2008).

Although online asynchronous discussions increase the flexibility of time, the anxiety about time is not eliminated. The diminution of timely responses, loss of a logical sequence of postings, tight time limit of discussions, and the varied time preferences of participation among different users have all produced the problem of arrhythmia. The existing literature tends to address these challenges by better catering for the temporal preferences of individual participants. Admittedly, the sensitivity and experience of time are of the individuals' own consciousness and control. However, the constraint of time cannot be overcome without knowing how the collective react to temporal patterns, expectations, and rules over time (Walther and Tidwell, 1995). A fair amount of research has concerned how individual posts or participants' features shape the amount of waiting time (Hewitt and Teplovs, 1999; Jeong, 2004; Tatsumi and Nakazawa, 2017). However, most of these studies are cross-sectional and regard the characteristics of posts or participants as being time-constant. This article tries to introduce a time-varying dimension to our understanding of online discussions, by integrating our current knowledge on time as latency, sequence, density, and frequency.

Individual posts versus environment factors in online policy discussions

Citizen participation in online policy discussions encourages the exchange of ideas on policy issues and enables participants to build on each other's ideas as well as examine the rationales behind their thoughts (Lee and Kwak, 2011; Phang et al., 2014). The quantity and quality of policy discussions are important for both stimulating citizens' political efficacy and formulating effective public programs and inclusive policies (Irvin and Stansbury, 2004; Neshkova and Guo, 2012). During the process of online policy discussions, a high quality of contributions implies a wider variety of productive ideas for the formulation of better policies. Moreover, a large number of posts and replies may attract more people to the discussion space and create a self-sustaining force for continuing engagement (Phang et al., 2014). However, such active and quality participation and interactions cannot be assumed but are influenced by a range of factors.

Scholars have identified several categories of factors that influence users' motivation to engage in online discussions. Content factors in the context of online policy discussions have focused a lot on the quality of individual posts. Although a longer waiting time is needed, high-quality posts that are posted early receive a greater overall response rate than other posts (Jeong, 2004; Jeong and Frazier, 2008; Tatsumi and Nakazawa, 2017): indicating that discussants seek to absorb the most valuable and constructive ideas within a limited time frame. High-quality posts often contain more information and

better justification, which invites viewers to comment on them. Chmiel et al. (2011a, 2011b) found that posts with negative emotions keep their vitality for a longer period of time and active users on the forum express predominantly negative sentiments. Negative emotions such as anger or anxiety function as warning signals that make us aware that things are not “all good,” thus urging us to react quickly (Clark and Fessler, 2015; Deonna and Teroni, 2017). Post opinion is another influential content factor. Studies (Chen and Chiu, 2008; Tatsumi and Nakazawa, 2017) have found that disagreeing views on the forum could elicit more responses, because we typically change our assessment of policies, or our behaviors, when something goes wrong, not when something goes right (Soroka and Mcadams, 2015). So when disagreeing views are expressed, they become catalysts for change and motivate other discussants to be responsive and contribute their ideas. Therefore, we propose the following hypotheses:

H1. Posts made in the early time phase on the forum have a higher possibility of getting a response than those made during the late phase.

H2. Posts with a negative sentiment have a higher possibility of getting a response than those with a positive sentiment.

H3. The higher the quality of a post is, the higher the possibility the post will get a response.

H4. Posts with disagreement have a higher possibility of getting a response than those with agreement.

The focus on individual posts has drawn our attention to the time-constant dimension of forum discussions, because when a post is made, the author rarely changes or revises it. However, online discussions are dynamic and change over time. One such time-varying type of factors is environment factors, such as discussion crowdedness. A post has to compete with many other posts to acquire the limited attention of the viewers. Moreover, the number of other posts at the time of posting influences the likelihood of this new post receiving a reply (Tatsumi and Nakazawa, 2017). If there are more other posts, it is more likely that a new post may be displayed for only a short period of time and therefore draw less attention in a discussion. Meanwhile, the number of active users (i.e. frequency) also varies according to time. Even if there are fewer posts at a given time, the presence of fewer users makes a new post less likely to be noticed (Tatsumi and Nakazawa, 2017). Therefore, the crowdedness of a forum is influenced by the concurrent factors of the density and the number of viewers. We thus propose the following hypothesis:

H5. Posts within a less crowded environment have a higher possibility of getting a response than those within a more crowded environment.

Another salient category of environment factors concerns the content environment, such as quality environment, sentiment environment, and opinion diversity. Few studies have transformed the quality and sentiment of an individual post into environment variables that

change over time, and even fewer studies have explored how these environment factors could moderate the impact of individual post factors. At this stage, we may expect that the effects of the quality and sentiment environment will function similarly as their time-constant counterparts. In other words, a higher quality environment and more negative sentiment environment may invite more responses. With regard to opinion diversity, scholars (Medaglia and Yang, 2017; Tatsumi and Nakazawa, 2017) found that the majority of threads responded to appear in an earlier divergence phase when opinions are diverse and the number of responses declines in the latter convergence phase when consensus is being formed. We can thus expect that higher opinion diversity will invite more responses. We thus propose the following hypotheses:

H6. Posts within a high-quality environment have a higher possibility of getting a response than those within a low-quality environment.

H7. Posts within a negative sentiment environment have a higher possibility of getting a response than those within a positive sentiment environment.

H8. Posts within a diverse opinion environment have a higher possibility of getting a response than those within a homogeneous opinion environment.

In addition, individual posts and environment factors could be interdependent on each other. Existing studies on this interaction between individual posts and environment factors have focused on the sentiment and opinion features. For instance, two studies (Medaglia and Yang, 2017; Tatsumi and Nakazawa, 2017) showed that people prefer to mingle with like-minded peers and reach a consensus as time goes by, leading to a lower possibility that a post with distinctive views will be replied to. Similarly, Chmiel et al. (2011a) found that the sentiment of a post is analogous to the sentiment environment shaped by the previous posts, and there are more no-response posts over time when someone expresses positive sentiments in an increasingly positive environment (Medaglia and Yang, 2017; Tatsumi and Nakazawa, 2017). However, most of the above findings are descriptive observations made by simply juxtaposing the evolving patterns of different variables. Whether these interactions are statistically significant is still unknown. Moreover, how the interactions between the individual post and environment features affect the response pattern is under exploration. An example is here: since participants need time to reflect and digest the high-quality posts (Jeong and Frazier, 2008), such a cognitive burden may inhibit one's will to reply to a new post, especially when the crowdedness or the overall quality of the posts is high. We thus ask the following:

RQ. How do the interactions between individual post factors and environment factors affect the possibility of getting a response?

In summary, existing research in the context of online policy discussions has focused very much on individual posts factors with static measures. In contrast to the psychological approach that focuses on participants' motivations to interact in an online discussion, we include both time-constant post variables (i.e. positing time, post quality, post

sentiment, and post opinion) and time-varying environment ones (i.e. crowdedness, quality environment, sentiment environment, and opinion diversity) in our explanations of why a post gets a reply. Moreover, we recognize the potential interaction between the individual posts and environment factors and try to provide a robust statistical verification of such interactions.

Methods

To measure how individual post and environment factors affect the responses, a multi-level event history analysis was used to examine the eight hypotheses and one research question. Event history data aim to explain whether and when events occur (De Nooy, 2011). Our data are suitable for conducting an event history analyses as we know the exact time at which participants receive a reply during the observation period. Compared with De Nooy's (2011) approach, our analysis is similar in the sense that we also took a longitudinal perspective by including time-varying variables and used multi-level analyses to accommodate the data structure (individual posts are nested in the environment). However, our analysis is different from De Nooy's in terms of the time-varying variables (i.e. level 2 variables). While De Nooy (2011) examined individual actors' characteristics that change over time (e.g. how many reviews one has written in the past 24 months), our time-varying variables center around the environment which is always evolving (e.g. the overall crowdedness of the forum).

When applying the event history analysis to online discussion data, we took 1 day as a unit for observation. The event to be explained was the occurrence of a new interaction, involving a pair of one post and one reply, for the first time. Those posts, which had received their first response, would be censored. In other words, we focused more on the expansion of the "scale" of the interactions rather than the accumulative "number" of interactions.

Since the data were collected at two levels, individual posts and the overall environment, we used multilevel analyses to assess the event data. This statistical tool is useful for studying nested and hierarchical data measured at multiple levels (Quintelier, 2010). Multilevel logistic models were used in this study because the dependent variable (DV) is a binary response (0=no response; 1=response), which means the relationship between the independent variables (IVs) and the DV is non-linear and the variance is heteroscedastic (Quintelier, 2010). The multi-level logistic equations are listed in the following:

Level-1 model (individual post layer)

$$\text{Prob}(Y=1|\beta) = P$$

$$\text{Log} \left[\frac{P}{(1-P)} \right] = B0 + B1*(SEN) + B2*(QUAL) + B3*(OPIN) + B4*(TIME)$$

Level-2 model (environment layer)

$$\begin{aligned}
B0 &= G00 + G01*(CROWDN) + G02*(DIVERS) + G03*(SENTEN) \\
&\quad + G04*(QUALIEN) + U0 \\
B1 &= G10 + G11*(SENTEN) + U1 \\
B2 &= G20 + G21*(CROWDN) + G22*(QUALIEN) + U2 \\
B3 &= G30 + G31*(DIVERS) + U3 \\
B4 &= G40
\end{aligned}$$

The level-1 model represents the effects of the explanatory variables on the response behavior at the individual post level, specifically, the main effects of time-constant individual level post variables. Among them, SEN indicates the sentiment of a post, QUAL is the quality of a post, OPIN consists of two dummy variables, which indicate whether a post contains disagreement or a neutral opinion, and TIME represents a dummy variable, indicating the early or late phase of the discussion. Since the time-varying variables are significantly correlated with posting time, the two groups of variables cannot be put into one model to satisfy the assumption of no multicollinearity. We thus removed the posting time when we put level-1 and level-2 variables into one model.

Next, environmental factors and random effects (U0–U3) are added to the level-2 model. In the level-2 model, while the equation B0 is used to test the effects of environment factors on the intercept of level-1 model, the equations B1–B3 are used to explore the potential interaction effects between post factors and level-2 predictors. The method of centering the level-1 predictors has a profound impact on the meaning of DVs, although there is no statistically correct choice (Enders and Tofighi, 2007). We decided to regard the level-1 predictors as the raw values, because in longitudinal multi-level models, the temporal factors at the microlevel usually center around the fixed values rather than the mean (Enders and Tofighi, 2007). In addition, TIME and OPIN are categorical variables, so their means are not meaningful. For level-2 variables, we selected the grand mean centering to reduce the impact of multicollinearity. The final models listed in Table 1 have all passed the collinearity (variance inflation factor (VIF) < 5) and overdispersion tests (residual deviance ratio < 1).

Participants and the platform

In order to systematically test the posting-response hypotheses, we designed an online policy discussion forum, Online Deliberation Singapore (ODSG for short), which was purposely established to meet the analytical requirements. We used an open source forum service, Vanilla Forums, to design our discussion platform. The core of the platform is written in PHP, and data are stored in a MySQL database. Given that Vanilla Forums accepts multiple plugins, we did tweak or develop new plugins to fit our research needs (see more details in Menon and Zhang, 2014; Perrault and Zhang, 2019). In the participant's interface, the platform appears to be a traditional tree-structure discussion forum, while contents of posts and replies, pseudo names of posters, and the posting time were the available information. All of the posts and replies were by default displayed in a chronological order.

Table 1. Estimated effects of individual and environment variables on receiving a response.

Dependent variables	Model 1		Model 2		Model 3	
	B	SE	B	SE	B	SE
Intercept	-1.73***	0.11	-2.55***	0.10	-2.46***	0.14
Time-constant post variables						
Posting time (reference: early phase)						
Late phase	-0.65***	0.11				
Sentiment	-0.44***	0.09	-0.53***	0.11	-0.37*	0.17
Quality	0.30***	0.07	0.36***	0.09	0.21*	0.10
Opinion (reference: agreement)						
Disagreement	0.68***	0.10	0.72***	0.12	0.73***	0.23
Neutral	0.88***	0.10	1.08***	0.13	1.03***	0.17
Time-varying environment variables						
Crowdedness			-0.08***	0.01	-0.06***	0.01
Sentiment environment (Se)			-6.71	3.96	-5.51	4.05
Quality environment (Qe)			9.08**	3.23	12.35***	3.74
Opinion diversity (Od)			15.84***	3.17	14.65***	4.50
Cross-level interactions						
Sentiment × Se					10.20*	5.12
Quality × Qe					-12.88*	4.71
Quality × crowdedness					-0.01	0.01
Opinion × Od (reference: agreement)						
Disagreement × Od					16.99**	5.50
Neutral × Od					-13.75*	5.92
N	4898		4894		4889	
df	1797		1793		1793	
Pseudo R-square	0.06		0.20		0.32	
Log likelihood	-2016.41		-1934.5		-1905.4	

SE: standard error; B: unstandardized coefficient.

* $p < .05$; ** $p < .01$; *** $p < .001$.

The operation of ODSG has gone through four phases: (1) the education phase, during which we drafted the issue education materials with the verification from policy makers, and developed them into easy-to-read interactive PowerPoint slides; (2) the recruitment phase, during which we sent invitations to citizens in Singapore through the access to an online panel provided by Yougov, an online survey company. Demographic quotas were set to match the most recent census data in Singapore and quota sampling was performed to finally get 510 citizens, who accepted our invitation and engaged in the next phase; (3) the discussion phase, during which the participants used our platform to discuss the issue of fertility. Trained moderators were present to eliminate bullying or trolling or identify spammers, but other than these tasks, the moderators did not directly intervene with the discussion process; (4) the reflection phase, during which we interviewed and surveyed

the users and policy makers to get their evaluations and feedback (Zhang and Soon, 2017; Perrault and Zhang, 2019).

In this article, we only present data relevant to our multi-level event history analysis from phase 3. We examined an 8-day discussion phase spanning from May 4 to 11, 2016. The discussion forum was locked after the 8 days lapsed, which means participants can still view the posts but cannot change/post any content. In total, 334 users logged in at least once and 145 users posted their opinions on the forum during the 8 days. A total of 743 out of 1798 posts on the forum received at least one reply during the observation period (i.e. 1055 no-response posts).

Measurements

The DV of the multi-level analysis is a binary variable, namely, whether or not a post gets a reply within the observed time period (1 = Yes, 0 = No). Detailed measurements of IVs are shown in the following.

Time-constant IVs

Posting time. Posting time refers to the day when a participant published a post in the forum, ranging from 1 to 8 (e.g. 2 means day 2). The variable was traced and recorded by the time stamps of actions on the forum. Posting time is a categorical variable, so we split it into two halves, the early phase (days 1–4) and the late phase (days 5–8). We regarded the early phase as the reference and compared the effect of later phase group with the reference. Overall, the highest post frequency occurred on day 5 and day 7, with 420 and 391 posts, respectively.

Quality. Cappella et al. (2002) argued that the level of argumentative complexity and depth of knowledge should be considered when judging the quality of participants' opinions. Based on their arguments, we evaluated the post quality according to its level of justification. We rated the post quality on a four-point scale (see Appendix 1). A group of six research assistants were trained to carry out the coding work, three of them were graduate students majoring in communications, while the other three were undergraduates majoring in political science. The three-round training process was run from 8 May to 21 May 2017 on a subset of 611 posts. After each round, the inter-coder reliability was calculated using ReCal, an online inter-coder reliability Web service that offers a variety of coefficients (<http://dfreelon.org/utills/recalfront/>). Based on the results of the reliability check, the team leader reconciled the disagreements in cooperation with the other five coders in face-to-face meetings. Items with low reliability were discussed in detail until an acceptable level of agreement was reached; examples of challenging posts were also presented and discussed in the meetings, so each coder had an idea of how to code such posts in the future. After the training process, we randomly selected 10% of the sample posts to test the inter-coder reliability (Krippendorff's $\alpha = 0.75$), which was satisfactory. The actual coding process took around 3 months, from May to July 2017. On average, the posts had a quality rating of 0.28 (standard error (SD) = 0.51).

Sentiment. Each message was classified ranging from -1 (negative) to 1 (positive) using AlchemyAPI (<https://www.ibm.com/watson/alchemy-api.html>), a service now

merged with IBM Watson. The API uses natural language processing technology and machine learning algorithms to extract semantic meta-data from the content, such as concepts, emotions, and sentiments. The application does not rely on bags of words, instead a Recurrent Neural Network was employed to dynamically capture sentiments (according to an interview here: <https://www.ibm.com/cloud/blog/cognitive-apis-with-watson-sentiment-analysis>). Alchemy's training data included tweets and Wikipedia articles. The application has been widely used in many research projects. A quick search on Google Scholar suggests that the term "Alchemy API" can be found in more than 3800 research papers. The application has evolved over the years and now offers various settings that can be changed, such as whether a message is evaluated holistically or according to its constituent parts. At the time of this study, the only option the application offered was its by-default holistic approach that resembles the continuous building of a neural network. On average, our posts had a sentiment rating of 0.09 ($SD=0.47$).

Opinion. All of the posts were first rated by their level of agreement with the policy expressed in the previous posts and separated into three groups: 1=disagreement (33.8%), 2=neutral (31.3%), and 3=agreement (34.9%). The same coding process used for measuring the opinion quality was followed here. The inter-coder reliability (Krippendorff's $\alpha=0.82$) for this measure is excellent. Since it is a categorical variable, we regarded the agreeing group as the reference and compared its effect with the neutral and disagreeing group when running our analyses.

Time-varying IVs

Crowdedness. The discussion crowdedness represents the extent of competition in the forum and the extent of redundancy of participants' attention. Crowdedness is the ratio between the number of other posts and the number of unique viewers. The number of other posts indicates the number of accumulative "competitors" on and before the day when the post got a reply. The number of unique visitors refers to the number of participants who logged into the forum on the day when the post got a reply. Both variables were traced and recorded by the time stamps of actions on the forum. Overall, the average of crowdedness in a period of 8 days was 16.15 ($SD=9.01$).

Temporal-quality. The quality environment of the day is the average quality points of all the posts on and before the day when a post received a reply. On average, the quality environment had a quality rating of 0.30 ($SD=0.02$).

Temporal-sentiment. The sentiment environment of the day is the average sentiment points of all the posts on and before the day when a post received a reply. On average, the sentiment environment had a rating of 0.08 ($SD=0.02$).

Temporal-opinion diversity. Stromer-Galley (2003) defined diversity as follows: "there are participants in the dialogue with distinct views on a particular issue." The diversity of opinions is the key factor affecting the cross-opinion interaction over time (e.g. Moy and Gastil, 2006). Diversity of opinions in this study was measured as the standard deviation among different views. We calculated the standard deviations among posts that belonged

to the three different agreement categories (1 = disagree, 2 = neutral, and 3 = agree) on and before the day when a post received a reply. The higher the standard deviation was, the more diverse the opinion environment would be regarded. On average, the diversity of opinion had a rating of 0.85 ($SD=0.02$).

Data analysis

We retrieved 1798 posts (including original posts and replies) and other time-series information in the forum using a spider software program specifically developed for this study. To test the hypotheses, a multi-level logistic regression analysis was conducted by *R*, using time-constant individual and time-varying environment variables as IVs, whereas getting a reply or not was the DV.

Results

Table 1 illustrates how the time-constant and time-varying variables are incorporated into the multilevel discrete-time event history model. Model 1 shows that in comparison with the posts posted in the early phase, the possibility of getting a reply for a post posted in the later phase is significantly lower ($B=-0.65, p<.001$). Thus, H1 is supported. The sentiment variable, which varies from -1 to 1 , reveals that posts with negative sentiments are more likely to get a reply than those with positive sentiments ($B=-0.44, p<.001$), which supports H2. In comparison with posts with low-quality, high-quality posts have a higher possibility of being replied to ($B=0.30, p<.001$), which supports H3. The opinion variable, which is categorized as disagreement, neutral, and agreement, reveals that posts with disagreeing ($B=0.67, p<.001$) and neutral views ($B=0.88, p<.001$) are more likely to get a reply than those with agreeing views. Therefore, H4 is supported.

The environment variables are added into Model 2. We found that the greater the crowdedness of the forum ($B=-0.06, p<.001$), the lower the possibility of getting a reply. When the number of posts increased on the forum and the competition for the scarce user attention increased, the possibility of getting a reply decreased. Therefore, H5 is supported. Conversely, the quality environment is found to significantly motivate response behaviors. In comparison with posts in a low-quality environment, posts in a high-quality environment have higher chances of being replied to ($B=9.08, p<.001$), which supports H6. A similar finding also supports the diversity hypothesis (H8). The more diverse the environment is, the higher the possibility that a post will get a reply ($B=15.84, p<.001$). The sentiment environment, however, does not show significant impacts on the occurrence of replies. Therefore, we rejected H7.

The results of cross-level interactions are shown in Model 3. The moderation effects of the environment variables on individual post variables are found to influence the slopes (B_1, B_2, B_3) in the level-1 equation. The results in Model 3 illustrate that the impact of post sentiment is significantly conditioned by sentiment environment ($B=10.20, p<.001$). The negative association between post sentiment and getting a reply is stronger for posts that exist in negative sentiment environment (see Figure 1). Conversely, the positive effect of high-quality posts on receiving a reply is inhibited by

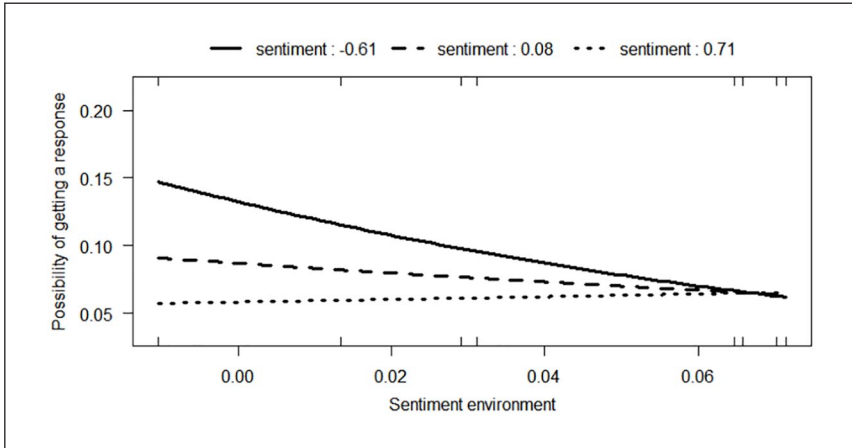


Figure 1. The interaction effect between the post sentiment and environment sentiment.

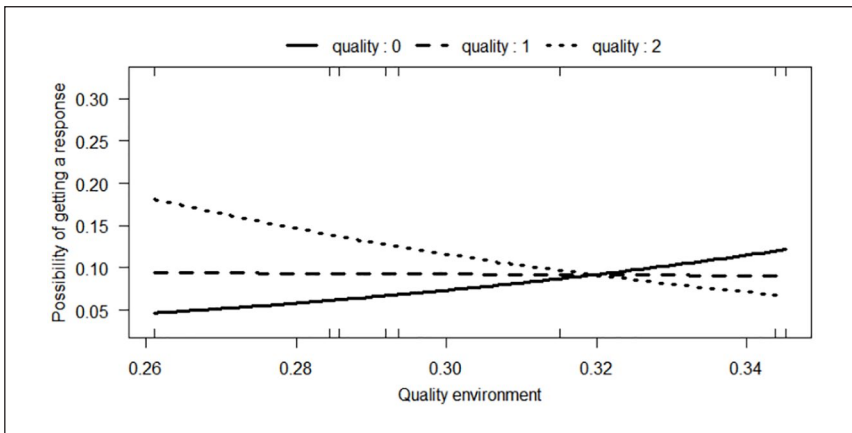


Figure 2. The interaction effect between the post quality and environment quality.

the high-quality environment ($B=-12.88, p < .01$) (see Figure 2). In terms of opinion diversity, it makes a difference to the impacts of posts with neutral ($B=-13.75, p < .05$) and disagreeing opinions ($B=16.99, p < .01$). When the discussion environment is heterogeneous, the gap between the effects of agreeing and neutral opinions is narrowed. Disagreeing views, however, show a much stronger impact than the agreeing counterpart (see Figure 3).

Discussion and conclusion

The time-series nature of this study enables us to explain the dynamics of interaction, defined as a post getting a reply, in online discussions over time. This is achieved by

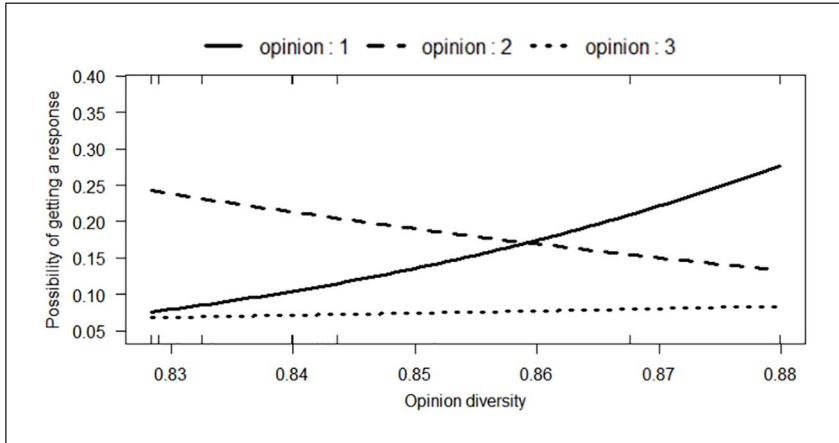


Figure 3. The interaction effect between the post opinion and environment diversity. The 10th, 50th, and 90th percentiles of sentiment variable values were employed in Figure 1.

testing a series of variables that are both under one participant's control (i.e. the individual post factors) and independent of one's control (i.e. the discussion environment during a specific time period). For the time-constant post factors, the results clearly indicate the advantages of early, emotionally negative, high-quality, and disagreeing posts on an online forum, consistent with the existing literature (Chen and Chiu, 2008; Clark and Fessler, 2015; Deonna and Teroni, 2017; Jeong and Frazier, 2008; Soraka and McAdams, 2015).

In addition to confirming the importance of individual post features, our findings further show that posts in a highly crowded environment were found to be less likely to get replies. Some previous studies (e.g. Jeong and Frazier, 2008; Tatsumi and Nakazawa, 2017) already regarded the high density of a forum as having a negative impact on getting replies. Other studies argued that interaction would only be possible when there is a large number of logged-in participants (Phang, Kankanhalli & Huang, 2014). Our study further combined the two factors, the number of posts and unique viewers, and we found that it is the proportion of logged-in users to the number of competing posts that really matters. If there are lots of participants, even a large number of posts may get noticed; if there are few participants, a small number of posts may be ignored. In contrast, the diverse and high-quality environment showed significantly positive impacts. As previous studies (e.g. Lee and Van Dolen, 2015; Stromer-Galley, 2003) stipulated that as the contested opinion climate alleviates the pressure to follow the majority, policy discussions are more likely to promote interactions if the discussion involves distinctive views rather than being indifferent, or having a high level of consensus.

Another set of interesting findings is that the environment factors had moderating effects on the relationship between individual post factors and getting a response. The quality environment moderated the relationship between the post quality and getting a response. Specifically, when the environment had low-quality, a high-quality post had a

much higher possibility of being replied to than a low-quality post. This finding suggests that a contrast effect occurred, with high-quality posts standing out and getting replies in a low-quality environment. However, we must note that the contrast effect was not that strong when the environment had high-quality: a low-quality post had a small advantage over a high-quality post. In other words, a contrast effect existed with regard to quality, but the size of the effect is smaller when the quality environment gets higher.

We also found that the sentiment environment had an indirect impact on response behavior by moderating the effect of post sentiment. Chmiel et al. (2011a, 2011b) emphasized the contagious, transitive, and compliant tendencies of sentiment expressions in online discussions. Our findings suggest an equalizing effect of a positive environment apart from the contagious effect. In particular, when the sentiment environment was negative, a negative post had a much higher possibility of being replied to than a positive post. A contagious effect (or the negative brew the negative) seemed to hold true here. However, the advantage of a negative post over a positive post decreased as the sentiment environment became more and more positive. When the environment sentiment reached its most positive state, all posts, regardless of their sentiment, had a similar possibility of being replied to. This suggests that a positive sentiment environment could equalize the reply possibilities of posts with different sentiments.

One final interesting finding lies in the interaction between post opinion and opinion diversity. For a single post, agreeing posts seemed to be the least powerful in eliciting responses. ~~However, this is only true when the opinion environment is more heterogeneous. When opinion diversity is low, agreeing posts actually have higher possibility of being replied to, compared with disagreeing and neutral posts.~~ Echoing Medaglia and Yang's (2017) finding that cross-minded interactions are more frequent in a diverse opinion climate, we noticed that disagreeing posts indeed invite more replies only when the opinion environment is more diverse. A hidden norm seems to manifest in the heterogeneous context: high diversity could boost the occurrence of new interactions and the growth of the thread, but only if one clearly disagrees with another existing stance in his/her post.

Through examining both the time-constant individual post level and the time-varying environment level variables, our research further extends our understanding about the role of time. Previous studies have tried to explain a post not getting a reply from the participants' perspective, such as limited time, distraction, or lack of interest in the topic (Dringus and Ellis, 2010; Jeong, 2004). Time has mostly been regarded as a resource that shapes people's consciousness and their activities (Meyer, 2003). The disorder and inconsistency between multiple ways of using time is believed to be addressed by adjusting individual participants' ability and motivation to allocate their precious time resources. The discussion forum, as a space linking polyrhythmia that facilitates a communal understanding of time use in Lefebvre's (2004) mind, seems not to function well.

We believe that time is a changing scene that is constantly shaped by each online behavior from all participants. If the existing studies had already alerted us about the social shaping of time (Hall, 1983; Walther and Tidwell, 1995), our study reveals that there is not one single unitary and fixed temporal norm or "primary-level culture" in online discussion forums. In other words, the temporal structuring of the asynchronous discussion relies on the synchronization of time practices that become part of how "I"

and “others” collaborate and get things done, echoing what Lefebvre (2004) referred to as multiple rhythms. If participants want their views to be replied to by others, they should not only conduct their online behavior based on their own preferences, but also need to consider the ever-changing contextual norms and others’ preferences in a specific time scene. Discussion organizers should identify what types of topics or what kinds of discussions are most appropriate at a specific time point. For instance, to counter the downward trend of getting a reply in a low-quality environment, moderators can call for high-quality contributions to push the discussions forward.

Limitations and implications

This article has a few limitations. First, previous scholars have traced the online discussion behaviors for a month or longer (Medaglia and Yang, 2017; Tatsumi and Nakazawa, 2017), and we only gave our participants 8 days to discuss the issue of fertility in Singapore, which may not fully illustrate how users’ response patterns change over a longer period of time. Second, the topic under discussion is another influential variable. Whether the topic is obtrusive and controversial can affect the participants’ willingness to join the discussion (Becker et al., 2010). Future research could use different topics to test whether the findings in this study can be replicated. Finally, our examination was performed on a conventional online policy forum constructed for citizen users. However, public forums and comment sections today predominantly employ curatorial techniques, such as direct-reply sequencing, voting mechanisms, and algorithmic prioritization based on user profiles or the linguistic characteristics of the posts themselves. How these emerging technical platforms overcome or can still be beset by the time-related factors needs exploration.

Despite the limitations, this article does offer some insights into how we can understand the temporal norms of online asynchronous discussions. Taking time as an evolving scene, and considering the discussion forum as a space housing the interactions among various time practices, we argue that trying to adjust individuals’ time preferences is not the only solution to address the tension between time and participation. Moreover, we need to examine how multiple time use patterns to interact with each other and reflect on the issues of “when,” “what,” and “how” to build the environment for a productive online discussion. Such reflection is urgently needed as new platforms such as social media are constantly updating the temporal scene.

Our study on a traditional discussion forum helps us to understand social media, which is currently the most popular space for online discussions. As social media have a higher requirement for timely response than forums, researchers need to employ shorter time intervals as the unit of observation. This may help them better capture the unique temporal norms of social media. Furthermore, user practices on social media often face the problem of “context collapse” (Brandtzaeg and Lüders, 2018). Social media take posts out of their original logical or chronological information flow and sort the sequence of contents (e.g. Facebook’s timeline) based on one’s interests and social network behaviors (Pond, 2016). On one hand, such an intervention from a platform design can make environment factors related to the chronological order (e.g. earlier posts are more likely to get replied) less predictive than content-related factors (e.g. sentiment environment) in

terms of eliciting responses. On the other hand, we should treat algorithmic recommendation as another type of rhythm, or a temporal norm being introduced and reinforced by platform runners. Future studies should aim to better understand these algorithms and the temporal norms hidden behind them. For example, network-related time-varying predictors (e.g. changes of views and likes) need to be included in order to discover the collective construction of the temporal scene on social media.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research is based on a research project funded by Singapore's Ministry of Education (MOE2013-T2-1-105).

ORCID iD

Weiyu Zhang  <https://orcid.org/0000-0003-2924-6517>

References

- Barber B (1984) *Strong Democracy: Participatory Politics for a New Age*. Berkeley, CA: University of California Press.
- Becker AB, Dalrymple KE, Brossard D, et al. (2010) Getting citizens involved: how controversial policy debates stimulate issue participation during a political campaign. *International Journal of Public Opinion Research* 22(2): 181–203.
- Berge Z (1997) Characteristics of online teaching in post-secondary, formal education. *Educational Technology* 37(3): 35–47.
- Brandtzaeg PB and Lüdgers M (2018) Time collapse in social media: extending the context collapse. *Social Media + Society* 4. DOI: 10.1177/2056305118763349
- Cappella NJ, Price V and Nir L (2002) Argument repertoire as a reliable and valid measure of opinion quality: electronic dialogue during campaign 2000. *Political Communication* 19(1): 73–93.
- Chen B, Resendes M, Chai CS, et al. (2017) Two tales of time: uncovering the significance of sequential patterns among contribution types in knowledge-building discourse. *Interactive Learning Environments* 25(2): 162–175.
- Chen G and Chiu MM (2008) Online discussion processes: effects of earlier messages' evaluations, knowledge content, social cues and personal information on later messages. *Computers & Education* 50(3): 678–692.
- Chmiel A, Sienkiewicz J, Thelwall M, et al. (2011a). Collective emotions online and their influence on community life. *PLoS ONE* 6(7): e22207.
- Chmiel A, Sobkowicz P, Sienkiewicz J, et al. (2011b). Negative emotions boost user activity at BBC forum. *Physica A Statistical Mechanics & Its Applications* 390(16): 2936–2944.
- Choi S, Yang SW and Chen W (2018) Longitudinal change of an online political discussion forum: antecedents of discussion network size and evolution. *Journal of Computer-Mediated Communication* 23(5): 260–277.
- Clark JA and Fessler DMT (2015) The role of disgust in norms, and of norms in disgust research: why liberals shouldn't be morally disgusted by moral disgust. *Topoi* 34(2): 483–498.
- De Nooy W (2011) Networks of action and events over time. A multilevel discrete-time event history model for longitudinal network data. *Social Networks* 33: 31–40.

- Deonna JA and Teroni F (2017) Getting bodily feelings into emotional experience in the right way. *Emotion Review* 9(1): 55–63.
- Dringus LP and Ellis T (2010) Temporal transitions in participation flow in an asynchronous discussion forum. *Computers & Education* 54(2): 340–349.
- Durrington VA and Yu C (2004) It's the same only different: the effect the discussion moderator has on student participation in online class discussions. *Quarterly Review of Distance Education* 5(2): 89–100.
- Enders CK and Tofighi D (2007) Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods* 12(2): 121–138.
- Friess D and Eilders C (2015) A systematic review of online deliberation research. *Policy & Internet* 7(3): 319–339.
- Gibbs W, Simpson LD and Bernas RS (2008) An analysis of temporal norms in online discussions. *International Journal of Instructional Media* 35(1): 63–75.
- Hall ET (1983) *The Dance of Life: The Other Dimension of Time*. New York: Anchor Books.
- Hesse BW, Werner CM and Altman I (1988) Temporal aspects of computer-mediated communication. *Computers in Human Behavior* 4(2): 147–165.
- Hewitt J (2005) Toward an understanding of how threads die in asynchronous computer conferences. *Journal of the Learning Sciences* 14(4): 567–589.
- Hewitt J and Teplovs C (1999) An analysis of growth patterns in computer conferencing threads. In: *Proceedings of the 1999 conference on computer support for collaborative learning*. Palo Alto, CA: Stanford University. Available at: <https://dl.acm.org/citation.cfm?id=1150240>
- Howlett M and Goetz KH (2014) Introduction: time, temporality and timescapes in administration and policy. *International Review of Administrative Sciences* 80(3): 477–492.
- Irvin RA and Stansbury J (2004) Citizen participation in decision making: is it worth the effort? *Public Administration Review* 64(1): 55–65.
- Jeong AC (2004) The combined effects of response time and message content on growth patterns of discussion threads in computer-supported collaborative argumentation. *Journal of Distance Education* 19(1): 36–53.
- Jeong AC and Frazier S (2008) How day of posting affects level of critical discourse in asynchronous discussions and computer-supported collaborative argumentation. *British Journal of Educational Technology* 39(5): 875–887.
- Lee G and Kwak YH (2011) *An Open Government Implementation Model: Moving to Increased Public Engagement*. Washington, DC: IBM Center for the Business of Government. Available at: <http://www.businessofgovernment.org/report/open-government-implementation-model-moving-increased-public-engagement>
- Lee HHM and Van Dolen W (2015) Creative participation: collective sentiment in online co-creation communities. *Information & Management* 52(8): 951–964.
- Lefebvre H (2004) *Rhythmanalysis: Space, Time and Everyday Life* (trans. S Elden and G Moore). London: Continuum.
- Macfadyen LP and Dawson S (2010) Mining LMS data to develop an “early warning system” for educators: a proof of concept. *Computers & Education* 54(2): 588–599.
- Majchrzak A and Malhotra A (2016) Effect of knowledge-sharing trajectories on innovative outcomes in temporary online crowds. *Information Systems Research* 27(4): 685–703.
- Medaglia R and Yang Y (2017) Online public deliberation in china: evolution of interaction patterns and network homophily in the tianya discussion forum. *Information, Communication & Society* 20(5): 733–753.
- Menon S and Zhang W (2014) A value-based design approach for online deliberation. In: *Proceedings of the 2014 conference for E-democracy and open government Asia*, 4 December, pp. 81–96.

- Meyer KA (2003) Face-to-face versus threaded discussions: the role of time and higher-order thinking. *Journal of Asynchronous Learning Networks* 7(3): 55–65.
- Moy P and Gastil J (2006) Predicting deliberative conversation: the impact of discussion networks, media use, and political cognitions. *Political Communication* 23(4): 443–460.
- Neshkova M and Guo D (2012) Public participation and organizational performance: evidence from state agencies. *Journal of Public Administration Research and Theory* 22(2): 267–288.
- Perrault ST and Zhang W (2019) Effects of moderation and opinion heterogeneity on attitude towards the online deliberation experience. In: *Proceedings of the 2019 CHI conference on human factors in computing systems*, (paper no.17), 18 April 2019. ACM.
- Phang CW, Kankanhalli A and Huang L (2014) Drivers of quantity and quality of participation in online policy deliberation forums. *Journal of Management Information Systems* 31(3): 172–212.
- Pond P (2016) Twitter time: a temporal analysis of tweet streams during televised political debate. *Television & New Media* 17(2): 142–158.
- Quintelier E (2010) The effect of schools on political participation: a multilevel logistic analysis. *Research Papers in Education* 25(2): 137–154.
- Soroka S and McAdams S (2015) News, politics, and negativity. *Political Communication* 32(1): 1–22.
- Stromer-Galley J (2003) Diversity of political conversation on the internet: users' perspectives. *Journal of Computer-Mediated Communication* 8(3): JCMC836. Available at: <https://academic.oup.com/jcmc/article/8/3/JCMC836/4584286>
- Tatsumi T and Nakazawa T (2017) Why does no one respond? An analysis of threads with no responses in online asynchronous discussions. In: *International conference on deliberation and decision making*, Singapore, 23–24 June.
- Walther JB and Tidwell LC (1995) Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. *Journal of Organizational Computing* 5(4): 355–378.
- Zhang W, Cao X and Tran MN (2013) The structural features and the deliberative quality of online discussions. *Telematics and Informatics* 30(2): 74–86.
- Zhang W and Soon C (2017) The main and spillover effects of online deliberation: changes in opinions, informational media use and political activities. In: *Proceedings of the 2017 conference for E-democracy and open government (CeDEM)*, 17 May 2017, pp. 217–226. IEEE.

Author biographies

Weiyu Zhang is an associate professor at the Department of Communication and New Media, National University of Singapore. Her research focuses on civic engagement and Information and Communication Technologies, with an emphasis on Asia. Her current project is to develop and examine an online platform for citizen deliberation.

Yipeng Xi is a PhD student in the Department of Communications and New Media, National University of Singapore. His research focuses on political communication in authoritarian contexts. Currently, his works on civic engagement, public opinion and collective action have appeared in several Chinese and English journals.

Anfan Chen is a postdoctoral researcher at the Department of Science Communication and Science Education, University of Science and Technology of China. His research field lies in Computational Communication and Science Communication.

Appendix I

Code book for post quality and opinion variables

Post quality. To what extent a speech gives complete justification for demands. Complete justification contains a conclusion, a reason, and a linkage between conclusion and reason.

No justification. A speaker only says that X should or should not be done, but no reason is given.

Inferior justification. Here, a reason Y is given as to why X should or should not be done, but no linkage is made between X and Y—the inference is incomplete. This code also applies if a conclusion is merely supported with illustrations. When it provides claim and reason while applied context is missing.

Qualified justification. A linkage is made as to why one should expect that X contributes to or detracts from Y. A single such complete inference already qualifies for code 2. Besides claim and reason, applied context is present ($x < y < z$).

Sophisticated justification. Here at least two complete justifications are given, either two complete justifications for the same demand or complete justifications for two different demands. More than one instance.

Post opinion

Disagreement. Mere opposition for the comments of a prior poster. Use linguistic cues, such as “I disagree,” “I oppose,” “I don’t have the same view as ABC,” and so on.

Neutrality. A combination of opposition and support for the comments of a prior poster, or no clear linguistic signals of support or opposition on the views of prior posters.

Agreement. Mere support for the comments of a prior poster. Use linguistic cues, such as “I agree,” “I support,” “I have the same view as ABC,” and so on.