Online Communities' Sensitivity to Unexpected Information: The Roles of Governance and "Cross-talk"

Haomin Lin

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Arts

University of Washington

2025

Committee:

Benjamin Mako Hill

Wang Liao

Program Authorized to Offer Degree:
Communication

University of Washington

Abstract

Online Communities' Sensitivity to Unexpected Information: The Roles of Governance and "Cross-talk"

Haomin Lin

Chair of the Supervisory Committee:
Benjamin Mako Hill
Communication

To stay on-topic, online communities exclude unexpected information—content they have little or no prior knowledge of—whose on-topicness is uncertain. However, unexpected information is valuable because it can enhance community members' understanding of the information and drive online communities to explore novel topics that enrich their discussions. Of course, online communities vary in their sensitivity to unexpected information. Utilizing a dataset comprising discussions in 3,238 active communities on Reddit, this study investigates the effects of online communities' formal and informal structures on communities' sensitivity to unexpected information. The results indicate a negative association between formal structures represented by community governance and sensitivity to unexpected information, while showing a generally positive relationship between informal structures represented by "cross-talk" (i.e., conversations within a discussion that exclude the initiator of the discussion) and sensitivity. The findings uncover how cross-talk and community governance correlate with online communities' sensitivity to unexpected information, illuminating the effects of the informal and formal structures of online communities.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION & BACKGROUND

Unlike traditional public spheres (Ferree and Gamson, 2002; Habermas, 1962), online communities frequently insist that discussions stay on-topic in ways that lead to high topic homogeneity (Benkler et al., 2015; Duguay, 2022). As a result, online communities reject discussions on unexpected information, i.e., content they have little prior knowledge of, because of uncertainty about the information's relevance to communities' conventional topics. However, unexpected information is valuable because it can raise people's cognitive attention to the information (Burgoon and Jones, 1976; Burgoon, 2015) and motivate systematic information processing, enabling people to obtain a deep understanding of the information (Chaiken and Ledgerwood, 2012). Besides, allowing discussions on unexpected information prepares online communities to develop novel discussions under their conventional topics, which enriches existing discussions and prevents communities from becoming outdated. Thus, the inclusion of unexpected information is meaningful for both individual members and online communities.

That said, online communities vary considerably in their inclusion of discussions on unexpected information, which I define as online communities' sensitivity to unexpected information. Why are some online communities sensitive to unexpected information while most are not? This study focuses on the influence of online communities' formal and informal structures. Online communities, organized by both formal and informal structures, develop their own institutions (Meyer and Rowan, 1977) that influence members' behaviors. These institutions result in varied sensitivity to unexpected information at the community level. This study focuses on examining the effects of two aspects of institutions of an online community—governance (representing formal structures) and 'cross-talk' (i.e., conversations within a discussion that do not involve the host or initiator of the discussion, representing informal structures)—considering their particular relevance to sensitivity.

The relevance of governance and cross-talk to sensitivity lies in the unexpectedness of information. In this study, unexpected information is characterized by presence of two factors: unprecedentedness and off-topicness. I use the unprecedentness to describe information that is new or that has not appeared in a context. For example, the term "COVID" was not known to almost anyone before the WHO made it the name for the disease that caused the 2019 Novel Coronavirus outbreak. Such unprecedentedness induces uncertainty for members when a new concept first emerges in an online community. I use the term 'off-topicness' to refer to a broad class of information that falls outside the normative scope of a community's discussions. For example, COVID would likely be off-topic in a Harry Potter fan community, although it might be on-topic—even though it is unprecedented—in a community for discussing respiratory viruses. Similarly, information might be expected even if they are off-topic (e.g., repeated spam messages). Unexpected information refers to information that is both unprecedented and off-topic. As governance and cross-talk form institutions that shape how community members address unexpected information, these two aspects of communities' institutions are particularly relevant to sensitivity.

Specifically, online communities typically govern themselves by moderating discussions (Kiesler et al., 2012), which constitute one of their most salient formal structures (i.e., the fixed set of rules, procedures, and programs that organizations employ to coordinate and control activities as per Roethlisberger and Dickson, 1939). Many governance practices, represented by moderation, are meant to keep discussions on-topic within an online community (Grimmelmann, 2015), providing guidelines for members' behaviors. Stronger governance contributes to formal structures that direct community members to maintain the on-topicness of the community, thereby more likely to avoid discussion on potentially off-topic unexpected information.

At the same time, cross-talk is an important aspect of a community's informal structures (i.e., emergent patterns of interactions among organization members; Smith-Doerr and Powell, 2005). Within online communities, I use cross-talk to refer to conversations that sideline the author of the original post. Cross-talk undermines an online community's solidarity, i.e., individuals' contributions of private resources to their group's collective goods as per Hechter (1987). Specifically, cross-talk tends to organize individuals whose interests are more aligned and different from their community's conventional topics as subgroups (Op 'T Roodt et al., 2021). Individuals may contribute more to these subgroups' collective goods, i.e., on-topic discussions, instead of those of the larger community. Also, cross-talk reduces members' dependence on their communities for accessing discussions of their interests. Consequently, more participation in cross-talk decreases members' obligations to maintain their communities' on-topicness, which promotes an online community's sensitivity.

This study tests the effects of governance and cross-talk on online communities' sensitivity to unexpected information, utilizing a dataset of online discussions from 3,238 active subreddits gathered from Reddit following two public affairs in 2020 (the COVID-19 pandemic and the murder of George Floyd). This study measures the degrees of governance and cross-talk in each subreddit and conducts regression analyses to uncover their effects on the sensitivity. Because sensitivity is subject to the extent to which unexpected information is aligned with a subreddit's conventional topics, this study also investigates the interaction effects between such topic alignment and the two kinds of community structures. Part of the results indicate that stronger governance consistently reduces communities' sensitivity to unexpected information. However, this effect weakens when the unexpected information becomes more closely aligned with a subreddit's conventional topics. Meanwhile, this study finds partial support for cross-talk's positive association with sensitivity. This positive relationship also diminishes as topic alignment rises. Furthermore, when topic alignment reaches a relatively high level, increased cross-talk can even reduce sensitivity. Overall, our findings reveal the negative association between governance and online communities' sensitivity to unexpected information, as well as the generally positive association between cross-talk and sensitivity. These findings point to discussions on organizing behaviors that give rise to institutions within online communities.

In the next section of this paper, I will review institutions' role in shaping sensitivity and the effects of formal and informal structures in online communities with a focus on governance and cross-talk, respectively. Along with the review, I will propose two sets of hypotheses regarding the effects of governance and cross-talk on online communities' sensitivity to unexpected information. Next, this paper will introduce the methods implemented in this study and report the results. Findings will be discussed lastly.

1.1 Institutions and Online Communities' Sensitivity to Unexpected Information

Institutions are constraints that structure human interactions (North, 1991). Within online communities, the effectiveness of institutions depends on the micro-level actions of individuals, instead of institutionalized products or policies, such as the protocols of moderation. Directed by mimetic isomorphism resulting from the standard response to uncertainty (DiMaggio and Powell, 1983), people tend to adopt existing institutions that have been proven to be successful. For online communities, on-topicness is a salient outcome of such institutions (Reddy and Chandrasekharan, 2023). However, the adoption of institutions that lead to success is merely ceremonial and aims to attain legitimacy for online communities. It does not ensure that activities are aligned with the adopted institutions (Meyer and Rowan, 1977). Instead, according to Coleman (1990), macro-level factors result from the aggregate effect of individual behaviors. Thus, whether online communities' on-topicness regulates their members' behaviors and inhibits communities' sensitivity to unexpected information relies on the activities of community members.

Specifically, the activities of online community members give rise to the formal and informal structures that shape the effectiveness of institutions. Formal structures in organizations are typically defined by written rules, regulations, etc. that are intentionally designed to govern the behaviors of individuals within organizations (Adler and Borys, 1996). They are usually perceived to be most effective in realizing the goal of institutions because of the direct outcomes enabled by the coercive enforcement. As discussed above, the effectiveness of formal structures in governing is primarily relevant to the existence of human activities that ensure the implementation of formal structures. For example, online communities employ moderators to regulate their members' discussions by community policies, such as disallowing toxic language use (Papacharissi, 2004). Moderators may perform differently because of their varying available time and energy or their recognition of the requirement for their work (Kuo et al., 2023). These variations influence whether moderators in an online community can actively respond to messages that violate policies in the community, which can result in different formal structures. As a result, different formal structures induce varied institutions, leading to different sensitivity to unexpected information across online communities.

In contrast, informal structures consist of norms, shared understanding, etc., which include unwritten interaction patterns within organizations (Granovetter, 1985). Different interaction patterns can induce varied person-to-person and person-to-organization relationships (Lawler et al., 2009), which can have implications for how individuals behave under the influence of their organizations' institutions. For example, informal structures can regulate individuals' selection of whom they reply to, which can further influence the content of their conversations (Prell et al., 2010). With the habituation of these behaviors, the dependence of community members on their communities can change. Such a change influences members' solidarity with their communities. The varying solidarity provides different conditions for realizing the institutions adopted by communities, such as on-topicness. Since informal structures usually emerge from the habituation of micro-level interactions (Berger and Luckmann, 1966), how people interact at the individual level is the micro-level foundation of informal structures within organizations. In the following sections, I will examine how the formal and informal structures induced by community members' micro-level behaviors influence online communities' sensitivity to unexpected information with a focus on governance and cross-talk, respectively.

1.2 Governance and Online Communities' Sensitivity to Unexpected Information

As one of the most salient constituents of online communities' formal structures, governance has a straightforward influence on shaping communities' varied sensitivity to unexpected information. Content moderators' removal of undesired messages and communication with community members are an important part of governance in online communities (Grimmelmann, 2015; Reddy and Chandrasekharan, 2023). Within online communities, off-topic discussions, trolls, and incivility often pollute discussions and lead to the loss of online community members (Fiesler et al., 2018). Since the success of online communities focuses on attracting active members (Malinen, 2015), for which the quality of discussions is a critical factor, communities usually implement governance to regulate members' use of such detrimental elements in discussions. By removing off-topic discussions, successful governance contributes to the success of online communities and shapes their sensitivity to unexpected information.

Governance directly impacts online communities' sensitivity to unexpected information by modifying the content of discussions. Whenever someone posts a piece of unexpected information that is outside of the community's conventional topics in discussions, volunteer moderators can remove it as long as they locate the message and deem it to be violating community rules. The removal directly eliminates discussions on unexpected information and reduces online communities' sensitivity to unexpected information at face value.

In addition, the products of governance minimize the exposure to unexpected information for community members, lowering the chance that subsequent discussions mention the unexpected information and fostering coherent discourse around communities' shared interests. This topic coherence maintains the conventional topics of interest within the online community, guiding both current and prospective members to focus on relevant discussions and avoid introducing unexpected information, thus diminishing the community's sensitivity to it. Specifically, for current members of the online community, the coherence of conventional topics increases their perceived risks of being left out of discussions when they attempt to mention unexpected information (Schwämmlein and Wodzicki, 2012), motivating them to keep their contributions on-topic. On the other hand, prospective members of an online community often observe existing activities before they join (Gallagher and Savage, 2015). Consistent topics in existing discussions demonstrate to newcomers what the discussions within an online community should concentrate on. Knowing the conventions of the community, prospective members can decide whether to join the community and contribute. If they decide to contribute, they tend to be well-prepared to align what they discuss with the community's conventions (Gallagher and Savage, 2015), which prevents the introduction of unexpected information.

Furthermore, governance regulates community members' discussions by reducing their motivations to introduce undesired content. Although moderation in online communities is often invisible, in some contexts, it leaves visible marks. For example, the removal by moderators on Reddit will leave a "[removed]" mark on the removed content while keeping the associated discussions intact. These marks communicate to community members that governance is active and warn members of the consequences of violating rules (Srinivasan et al., 2019). Having messages removed is an undesired outcome for people who participate in online discussions. The pervasiveness of governance within online communities can effectively warn community members to avoid such consequences by obeying the communities' conventions of being on-topic. As a result, online communities' sensitivity to unexpected information will be reduced.

By modifying discussion content, presenting community conventions, and influencing community members' motivations, governance maintains the on-topic conventions of online communities and reduces an online community's sensitivity to unexpected information. Governance intensity should be negatively associated with online communities' sensitivity to unexpected information. Thus, this study hypothesizes: *(H1a) Governance intensity is negatively associated with online communities' sensitivity to unexpected information.*

That being said, governance's effects on sensitivity are not identical for all online communities. One of governance's major goals is to maintain on-topicness in discussions within online communities. But unexpected information is not always off-topic for online communities. Some online communities may find unexpected information aligned with their conventional topics from the beginning and present higher sensitivity with stronger governance. Additionally, higher topic alignment may increase the effectiveness of governance. Thus, unexpected information's alignment with communities' conventional topics, i.e., topic alignment, may positively moderate the relationship between governance intensity and sensitivity. With such consideration, this study hypothesizes: *(H1b) Unexpected information's topic alignment moderates the relationship between communities' governance intensity and their sensitivity to unexpected information such that greater topic alignment corresponds to weaker negative association between governance and sensitivity or stronger positive association between governance and sensitivity.*

1.3 Cross-talk and Online Communities' Sensitivity to Unexpected Information

Cross-talk increases communities' sensitivity to unexpected information by shaping the solidarity of online community members to their communities, which influences the effectiveness of communities' institutions that favor on-topicness. Solidarity within a group is the average proportion of each member's contribution of their private resources to the collective goods of the group (Hechter, 1987). The most important collective good for an online community is on-topic discussions organized under separate discussions within online communities. In online communities that strive to be spaces for focused discussion free from irrelevant "noise" (Ren and Kraut, 2014), only on-topic discussions can serve as a collective good to be consumed by all members. Considering online community members' time, or bandwidth (Kollock and Smith, 1996), as their private resource, the average contribution of members' bandwidth to on-topic discussions within an online community is an online community's solidarity. Lower solidarity corresponds to fewer contributions to on-topic discussions, reducing the effectiveness of institutionalized on-topicness and allowing for higher sensitivity to unexpected information in online communities.

Cross-talk induces a direct negative impact on an online community's solidarity. Cross-talk happens when two participants within a discussion exclude the host or initiator of the discussion in their conversations. When a host starts a discussion in an online community, they often set up a topic as the agenda for the discussion, which is often relevant to the community's conventional topics. Also, the host usually has better awareness of the set topic and is thus more capable of keeping conversations between the host and other discussants on-topic (Fujimoto, 2016). In contrast, cross-talk may include off-topic, personal content due to the absence of the host. In this way, cross-talk does not contribute to the collective good, i.e., the on-topic deliberation, of the discussion group. Furthermore, cross-talk costs discussion participants' bandwidth, reducing contributions to on-topic discussions and introducing undesired content that may damage the collective good. Increased cross-talk hence decreases online community members' solidarity, resulting in lower on-topicness and higher sensitivity to unexpected information within online communities.

In addition, cross-talk mars solidarity indirectly by implicating the two decisive factors for solidarity. The realization of solidarity hinges on both members' dependence on the group and the control implemented by the group (Hechter, 1987). Cross-talk sets members free from the dependence on their community by creating subgroups. In online discussion groups, members do not have to leave when they are not satisfied with the current discussion. Instead, they can find others with the same dissatisfaction via cross-talk and form subgroups (Gilson et al., 2015; Op 'T Roodt et al., 2021). Because of the more aligned interests, they are more willing to contribute to the collective good, i.e., the discussions on their shared interests, of their subgroup, and thus present high solidarity to their subgroup. Meanwhile, their dependence on the large discussion group, as well as their solidarity with the large group, decreases (Jarman, 2005).

Moreover, individuals' solidarity with their subgroups also reduces their amenability to the larger group's control, especially in online communities. Larger discussion groups in online communities are usually under the regulation of controls from the communities' moderators. The concern about their content being removed drives community members to present high solidarity by contributing to on-topic discussions. With a more dedicated subgroup emerging from cross-talk, members can get away from such control by moving their conversations to other online communities or even to dyadic direct messages since smaller groups are easier to manage (Tausczik and Huang, 2019). Thus, cross-talk further inhibits solidarity and on-topicness within online communities, developing higher sensitivity to unexpected information. This study thereby hypothesizes: *(H2a) Cross-talk intensity is positively associated with online communities' sensitivity to unexpected information.*

Similarly, the positive relationship between cross-talk and sensitivity is subject to the moderating effects of unexpected information's topic alignment. When the unexpected information is aligned with communities' conventional topics, reduced solidarity caused by increased cross-talk lowers discussions on on-topic information and decreases sensitivity. Conversely, off-topic unexpected information may be discussed more when solidarity is lower because of higher participation in cross-talk. In this case, sensitivity is higher when cross-talk is more pervasive in the online community. Thus, this study hypothesizes: *(H2b) Unexpected information's topic alignment moderates the relationship between communities'* *cross-talk intensity and their sensitivity, such that greater topic alignment corresponds to a weaker positive association between cross-talk intensity and sensitivity or a stronger negative association between cross-talk intensity and sensitivity.*

Chapter 2

METHODS

2.1 Dataset

To examine the effects of formal and informal structures on online communities' sensitivity to unexpected information, this study utilizes online discussions from Reddit. Specifically, I utilize a Reddit database downloaded via *pushift.io*(Baumgartner et al., 2020), a website that redistributed data from Reddit and other social media platforms, which has been widely used in social research. The entire dataset contains 1,344,306,994 posts and 13,126,553,528 comments from June 2005 to June 2021. As one of the largest online platforms, Reddit hosts thousands of active online communities referred to as subreddits. Subreddits deploy various governance approaches and contain discussions that feature different interaction patterns. Subreddits are typically organized around topics. For example, r/snowboarding[1] mostly discusses snowboards or snowboarding experience. Members of subreddits thus usually share an expectation for discussions happening in their subreddits (Hwang and Foote, 2021; Lin et al., 2017), and may react differently to unexpected information.

Because this study focuses on online communities whose governance practices and interaction patterns can provide observable variations among them, the subreddits to be analyzed should contain discussion posts that contain a large number of messages on average. In doing this, this study conducts a data sampling process that ensures active discussions within subreddits. Before that, a time range is required to narrow down the analysis of subreddit activity. As discussed earlier, unexpected information features unprecedentedness. To ensure such unprecedentedness, this study constructs the dataset focusing on the information about two public events: (1) the WHO's announcement of the official name for the COVID-19 pandemic that happened on February 11th, 2020, with the unexpected information as the keyword "COVID;" (2) the murder of George Floyd that happened on May 25th, 2020,

[1] Reddit prepend r/ to subreddits

with the unexpected information as the key phrase "George Floyd." The word "COVID" and the phrase "George Floyd" barely appeared in previous discussions before the events happened, ensuring the unprecedentedness of the information. Given that both events happened in 2020, this study focuses its analysis of subreddits' activeness on the 24 months from January 2019 to December 2020.

Using this time range, this study conducts a three-step data sampling process from the Reddit database. First, to be certain that all subreddits were continuously active across the 24 months, this study identifies all the subreddits that have both posts and comments in every month from January 2019 to December 2020 within the database (a total of 33,212 communities). Then, this study measures the activeness of each subreddit by the average number of comments per post over the 24 months and selects the top 10% most active subreddits, which narrows the pool to 3,322 subreddits. Lastly, because there is a concern that non-English texts may introduce noise to computing topic-based control variables to be discussed below, this study removes non-English subreddits. Specifically, this study concatenates 10 randomly sampled messages (i.e., posts or comments) from each subreddit and uses the langdetect package in Python to identify the languages of the joined sentence. 79 subreddits with non-English joined sample sentences are removed. This study also leaves out four subreddits that contain too many messages (r/askreddit, r/politics, r/unpopularopinion, and r/pics) to ensure computational efficiency. During the subsequent calculation of unexpected information' topic alignment, one subreddit is dropped for not being able to generate more than one top word for the calculation. The dataset ends up containing 3,238 subreddits.

2.2 Measures

2.2.1 Online Communities' Sensitivity to Unexpected Information

This study operationalizes online communities' sensitivity to unexpected information as the *proportion of messages mentioning unexpected information* within each online community in the four weeks after the corresponding public events happened (February 11th to March 10th, 2020, for "COVID", May 25th to June 22nd, 2020, for "George Floyd"). Because the sizes of online communities differ from one another, this study measures the *proportion of messages mentioning unexpected information*, i.e., including the corresponding keywords, in the community. Messages include both original posts and comments.

Specifically, this study implements the following three steps to measure sensitivity. To begin with, this study preprocesses the text data of online discussions by implementing regular expressions to remove all the URLs in the body of posts and comments, lowercasing all the texts, and joining all words into one string so that messages mentioning "George Floyd" using a hyphen or without the space can also be recorded. Secondly, this study retrieves all the processed strings that contain the case-insensitive string "covid" or the string "georgefloyd" as the messages that mention unexpected information. Finally, the number of messages that mention the related information is divided by the number of all the messages in the online community in the four weeks after the events. The proportions derived from the last step are the *proportion of messages mentioning unexpected information* that represents communities' sensitivity to unexpected information ($\mu = 0.0007$, $\sigma = 0.0022$, Median = 0, Min = 0, Max = 0.0525 for the COVID dataset, $\mu = 0.0010$, $\sigma = 0.0023$, Median = 0.00004, Min = 0, Max = 0.0462 for the George Floyd dataset).

2.2.2 Online Communities' Governance-Related Formal Structures

In this study, online communities' formal structures center on governance within online communities. This study employs three measurements of formal structures that reflect the extent to which an online community employs governance. Considering that voluntary moderation is a primary aspect of online community governance (Seering et al., 2019), the three measurements are all relevant to content moderation in online communities.

The first measurement directly measures the micro-level moderation activities in an online community. In online communities, content removal is an effective manifestation of community governance (Srinivasan et al., 2019; Weld et al., 2025). Therefore, the proportion of messages that were removed in an online community can reflect the intensity of moderation activities. This study measures the proportion of removed messages by dividing the number of messages whose body is "[removed]" by the total number of messages, i.e., the *removal proportion*, in each subreddit ($\mu = 0.0209$, $\sigma = 0.0337$, Median $= 0.0107$, Min $= 0$, Max $= 0.4081$ for the COVID dataset, $\mu = 0.0214$, $\sigma = 0.0339$, Median $= 0.0110$, Min $= 0$, Max $= 0.4500$ for the George Floyd dataset).

This study measures the *proportion of u/AutoModerator's*[2] *activities* among all messages in an online community by dividing the number of messages whose author is u/AutoModerator by the total number of messages in each subreddit ($\mu = 0.0061$, $\sigma = 0.0297$, Median $= 0$, Min $= 0$, Max $= 1$ for the COVID dataset, $\mu = 0.0062$, $\sigma = 0.0277$, Median $= 0$, Min $= 0$, Max $= 1$ for the George Floyd dataset). It serves as the second measurement of the online community's governance intensity. The u/Automoderator is a customized moderation bot that moderators set up to manage the posts and comments for a subreddit. The *proportion of u/Automoderator's activities* is associated with the practice of community governance. To make sure the measurement of governance is not susceptible to the impact of the corresponding unexpected information, the calculation of the above two variables uses data from the month that is three month earlier than the month of the related event (November 2019 for the "COVID" dataset, February 2020 for the "George Floyd" dataset).

Lastly, this study uses the *number of moderators* within each online community as another variable to gauge online communities' strength of governance. Using a publicly available dataset collected in 2024[3] that contains the lists of moderators for numerous subreddits, this study retrieves the *number of moderators* for each subreddit ($\mu = 6.2943$, $\sigma = 27.9295$, Median $= 4$, Min $= 0$, Max $= 1536$). Since moderators undertake the tasks of removing undesired content and communicating community policies to community members, the *number of moderators* represents an online community's competence in regulating its members' behaviors.

2.2.3 Online Communities' Cross-Talk-Related Informal Structures

This study's measurement of informal structures focuses on the extent to which cross-talk happens in an online community. To ensure the measurement of this informal structure is

[2] Reddit prepend u/ to users

[3] https://github.com/almayor/reddit-mods-dataset

comprehensive, this study computes three variables around cross-talk using data from the same months for calculating formal structures within online communities.

The first variable is the *average cross-talk proportion*, which measures an online community's average proportion of conversations that do not involve the original poster (OP) of a discussion post across all posts in an online community. This study operationalizes cross-talk on Reddit as the conversations that neither the receiver nor the sender is the OP of the discussion post. Measuring the proportion of messages that do not involve the OPs directly reflects the extent of cross-talk participation in a discussion post. To derive the patterns of cross-talk participation within an online community, this study computes the *average cross-talk proportion* by averaging the proportion of cross-talk conversations across discussion posts within an online community as the first measurement for online communities' informal structures ($\mu = 0.4095,\ \sigma = 0.1081,$ Median $= 0.4130$, Min $= 0$, Max $= 0.9850$ for the COVID dataset, $\mu = 0.4082,\ \sigma = 0.1068,$ Median $= 0.4113$, Min $= 0$, Max $= 0.9570$ for the George Floyd dataset).

The second measurement of informal structures is an online community's *average centralization* of discussion posts' directed interaction networks of message authors. In this network, each node represents an author, and a directed edge indicates a comment from the sender to the receiver. As shown in Equations (1) and (2), centralization is calculated as the ratio of the average difference between the highest degree centrality (C) and that of all other nodes to the difference between the highest and lowest degree centrality in the network (Shrum and Mullins, 1988). In the equations, N is the number of unique members in an interaction network, and $\alpha_{i,j}$ refers to an edge that connects author i and author j in the network.

$$C(Degree\ Centrality) = \frac{\sum_{j=1}^{N} \alpha_{i,j}}{N-1} \tag{2.1}$$

$$Centralization = \frac{\sum_{i=1}^{N} C_{max} - C_i}{N(C_{max} - C_{min})} \tag{2.2}$$

In discussion posts on Reddit, the OPs are more likely than other discussants to attract replies and center discussions around them. Higher centralization indicates more OP-others conversations and less cross-talk, and vice versa. Measuring centralization can evaluate the extent of cross-talk participation within a discussion post. To derive a variable that measures the community-level informal structures, this study computes the *average centralization* of an online community by averaging the centralization values of individual discussion posts as the second measurement. As shown in Equation (2), the calculation involves a fraction where the denominator is the product of the range of degree centrality and the total number of nodes in the network. The numerator is the sum of the differences between each node's degree centrality and the network's maximum degree centrality. The value of the fraction is the centralization of the network ($\mu = 0.7329,\ \sigma = 0.0695,\ \text{Median} = 0.7342$, $\text{Min} = 0$, $\text{Max} = 0.9943$ for the COVID dataset, $\mu = 0.7332,\ \sigma = 0.0666,\ \text{Median} = 0.7322$, $\text{Min} = 0$, $\text{Max} = 0.9954$ for the George Floyd dataset).

Another metric that measures informal structures is an online community's *average transitivity* of discussion posts' undirected interaction networks of message authors. As proposed by Newman et al. (2001) and applied by Dekker et al. (2019), equation (3) presents the formula for computing transitivity, whose value is decided by the number of triangles divided by the number of triads times three. Triangles refer to loops of length three where three edges connect three nodes, and each pair of nodes connects with each other. Triads are possible triangles where two edges share a node. As triangles indicate full connections among three users and each discussion post only has one OP, triangles within a post must include the connection between at least one pair of two community members, neither of whom is the OP. Such a connection is cross-talk. Thus, transitivity indicates the proportion of cross-talk participation in a discussion post. Similarly, this study averages the transitivity of each discussion post to compute the *average transitivity* for each community ($\mu = 0.07553,\ \sigma = 0.0411,\ \text{Median} = 0.0708$, $\text{Min} = 0$, $\text{Max} = 0.6449$ for the COVID dataset, $\mu = 0.0770,\ \sigma = 0.0402,\ \text{Median} = 0.0720$, $\text{Min} = 0$, $\text{Max} = 0.6835$ for the George Floyd dataset).

$$Transitivity = 3\frac{\#triangles}{\#triads} \tag{2.3}$$

2.2.4 Unexpected Information's Topic Alignment

As distinguished in the two hypotheses, this study factors in the moderating effects of unexpected information's alignment with an online community's conventional topics, i.e., topic alignment. To measure the topic alignment with the consideration of all the numerous texts within an online community, this study incorporates the TF-IDF method (Salton and Buckley, 1988) to conduct keyword extraction and word embeddings' cosine similarity to evaluate content similarity. This study extracts the five most important words from each subreddit to represent the conventional topics of the community.

Next, this study uses a BERT-based sentence transformer model to convert the key terms from subreddits into 384-dimensional numerical vectors. Finally, this study computes the cosine similarity between these vectors and vectors transformed from phrases representing unexpected information, i.e., unexpected information's *semantic similarity to topics*, as the measurement for topic alignment ($\mu = 0.1472$, $\sigma = 0.0272$, Median = 0.1469, Min = 0.0592, Max = 0.4064 for the COVID dataset, $\mu = 0.1271$, $\sigma = 0.0269$, Median = 0.1276, Min = 0.0303, Max = 0.3603 for the George Floyd dataset). Specifically, this study uses the phrases 'the covid pandemic' and 'the murder of george floyd' to provide adequate contexts for the model. The cosine similarity is calculated as the normalized dot product of the vectorized phrases and vectorized key terms. This study then averages the five key terms' similarity to each unexpected information separately as the final value of the corresponding unexpected information's *semantic similarity to topics* for different online communities. To ensure the model incorporates the contexts related to COVID and George Floyd into computation, this study uses the model 'all-mpnet-base-v2' trained on public datasets that contain content produced after 2020 (Siino, 2024).

2.2.5 Control Variables

Both network-based and content-based features of communities can influence how online communities organize discussions and implicate the behaviors of community members (Panzarasa et al., 2009). Specifically, the *size* and *volume* of online communities correspond to varied numbers of opportunities for discussions in an online community to include unexpected information. Besides, the density of discussion posts' interaction networks presents the connectedness among participants in discussions, introducing differing potential for random usage of unexpected information in discussions. To control for those effects, this study measures and controls the *size* ($\mu = 3873.4768,\ \sigma = 11720.9651,$ Median $= 921$, Min $= 1$, Max $= 227586$ for the COVID dataset, $\mu = 4011.3107,\ \sigma = 11969.1731,$ Median $= 996$, Min $= 1$, Max $= 246451$ for the George Floyd dataset), *volume* ($\mu = 19903.3789,\ \sigma = 66855.7501,$ Median $= 3971.5$, Min $= 5$, Max $= 1500117$ for the COVID dataset, $\mu = 20157.0608,\ \sigma = 62763.6477,$ Median $= 4,563$, Min $= 11$, Max $= 1239514$ for the George Floyd dataset), and *average density* ($\mu = 0.3306,\ \sigma = 0.8791,$ Median $= 0.2641$, Min $= 0$, Max $= 31.1000$ for the COVID dataset, $\mu = 0.3253,\ \sigma = 0.8653,$ Median $= 0.2694$, Min $= 0$, Max $= 31.3956$ for the George Floyd dataset) of interaction networks of an online community.

The *size* and *volume* of an online community refer to the number of unique community members and the number of messages in a subreddit, respectively. The density of an interaction network is calculated by dividing the number of actual edges by the maximum number of possible edges (Shrum and Mullins, 1988), as shown in Equation (7). E refers to the number of edges in an interaction network, while N denotes the number of unique users in the network. This study uses *average densities* of directed interaction networks of community members to derive a community-level metric that indicates the pattern of connectedness among community members within a subreddit.

$$Density = \frac{E}{N(N-1)} \tag{2.4}$$

On the other hand, the content of conventional discussions within online communities also implicates members' behaviors within subreddits. For example, discussions in subreddits focused on politics, sports, and current events tend to generate deeper and wider discussions (Yu et al., 2024), which may influence the chance of discussing unexpected information. To control for such effects, this study used the LDA (Latent Dirichlet Allocation) method (Blei et al., 2003) to derive topic distributions across online communities as the content-based control variables. Specifically, this study applies LDA topic modeling to texts sampled from the 3,238 selected subreddits. Each subreddit is represented by ten

500-word sample strings created by concatenating messages. This study determines the optimal number of topics (k) by selecting the k that can generate a model with the highest coherence score (Röder et al., 2015). The final model generated topic distributions for each sampled text, which are averaged to represent the subreddit's overall topic distribution. This process produces 120 topic variables for the COVID dataset and 15 topic variables for the George Floyd dataset. To avoid multicollinearity, this study removes one topic variable from each of the variable sets, making it 119 for the COVID dataset and 14 for the George Floyd dataset.

In order to maintain consistency in the temporality of the data, all the control variables are generated using data from the same months for calculating the independent variables.

2.3 Analytic Plan

To explore how formal and information structures within online communities relate to their responsiveness to unexpected information, this study employs two regression analyses, corresponding to the COVID dataset and the George Floyd dataset, respectively. Because the six independent variables are not strongly correlated with one another, the two regression models include all six independent variables. To make sure the regression analyses fit the features of the data, this study first inspects the dependent variable, *proportion of messages mentioning unexpected information*, and discovers that the data contains a large number of zero values because many subreddits did not react to the unexpected information within the four weeks after both relevant events. To accommodate this characteristic of the data, this study implements left-censored regression censored at zero as the regression method for data analysis, which can handle dependent variables that have a considerable number of zeros (Henningsen, 2010).

Furthermore, to meet the assumptions of regressions, this study inspects the distribution of dependent, independent, and control variables and implements log-transformation if it is necessary. As a result, *size*, *volume*, and *average density* of online communities are log-transformed. Also, this study standardizes the *number of moderators* before sending the corresponding variables to the regression models to assist the interpretation of regression results.

One last important note is that the variable *semantic similarity to topics* of unexpected information is not transformed or centered in the regression model. This means that the coefficients of the independent variables reflect these variables' marginal effects when the similarity equals zero (i.e., when the information is maximally off-topic). Since this study investigates the role of formal and informal structures in influencing the sensitivity to unexpected information, using highly off-topic information as the reference point aligns well with the off-topicness featured by unexpected information. Not centering the *semantic similarity to topics* is consistent with the study's conceptualization of unexpected information.

Chapter 3

RESULTS

3.1 Governance and Online Communities' Sensitivity to Unexpected Information

H1a predicts that governance in online communities is negatively associated with their sensitivity to unexpected information. As shown in Table 3.1, this hypothesis is partially supported by a statistically significant marginal effect on *removal proportion* and the *proportion of messages mentioning unexpected information* in the George Floyd dataset ($\beta = -0.2124$, $SE = 0.0451$, $p < 0.001$). It indicates that when the other variables are 0, changing a community from removing nothing to removing everything is associated with 21 fewer messages mentioning unexpected information every 100 messages. As a more realistic illustration of *removal proportion*'s effect, Figure 3.1 presents a non-linear association between the *removal proportion* and the predicted *proportion of messages mentioning unexpected information* when all the other variables take their median values. When the other variables are held at their median and the *removal proportion* is relatively low, a 1% increase in *removal proportion* from 1% to 2%, i.e., from removing one message to two messages every 100 messages, corresponds to a drop from 24 out of one 1000 messages mentioning the unexpected information to 13 out of 1000. In contrast, when a community moves from 4 messages to 5 messages getting removed every 100 messages, the *proportions of messages mentioning unexpected information* remain nearly unchanged and are close to zero. This suggests that the negative effect of *removal proportion* on the *proportion of messages mentioning unexpected information* is more pronounced at relatively lower levels of *removal proportion* and becomes negligible at higher levels (e.g., larger than 0.04). In the George Floyd dataset, 0.4 corresponds to the 86th percentile for *removal proportion.* This means that for most subreddits in the George Floyd dataset, strengthening their governance practice to remove 1 more message per 100 messages can considerably reduce the number of messages mentioning unexpected information. For a subreddit with a median volume of 4563, enhancing governance to move from removing 1 message to 2 messages per 100 messages results in 50 fewer messages mentioning unexpected information out of 1000 in the four weeks after the murder of George Floyd. In terms of adding to the exposure of unexpected information, this is a considerable change in sensitivity.

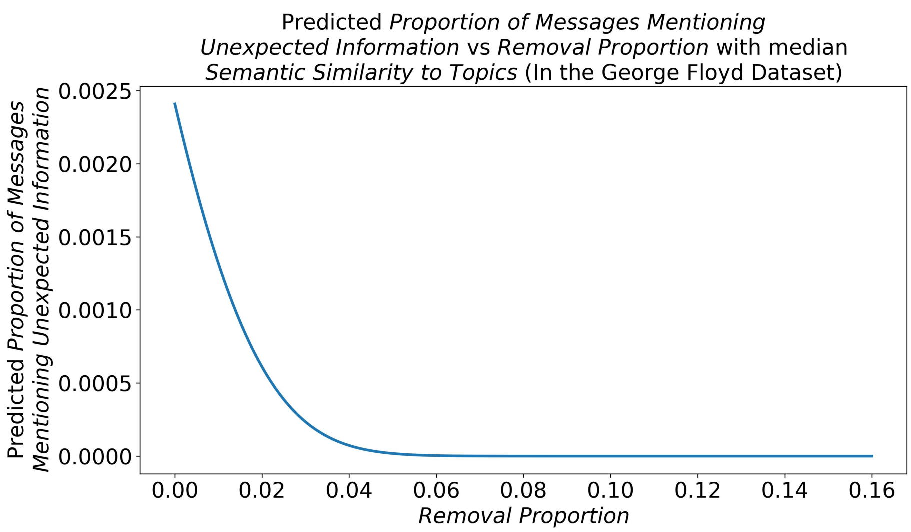

Figure 3.1: Predicted *proportion of messages mentioning unexpected information* as a function of *removal proportion* in the George Floyd dataset with unexpected information's *semantic similarity to topics* at its median. With median similarity, *removal proportion* has a negative relationship with the *proportion of messages mentioning unexpected information.* This relationship is stronger when *removal proportion* is relatively small (e.g., 0.1).

H1b predicts that topic alignment between unexpected information and a community's conventional topics moderates the relationship between governance and sensitivity with a positive interaction effect. Specifically, increased topic alignment would strengthen the positive effect of governance or weaken its negative effect. H1b is also partially supported. The results in Table 3.1 suggest a statistically significant relationship between the unexpected information's *semantic similarity to topics* and the marginal effect of *removal proportion* in the COVID dataset ($\beta = 0.1877,\ SE = 0.0904,\ p = 0.0379$). As the marginal effect of *removal proportion* is -0.2124, between a community that takes the information completely off-topic and another taking it completely on-topic, the effect of *removal proportion* a de-

Table 3.1: Censored regression results on how changes in variables of formal and informal structures affect the *proportion of messages mentioning unexpected information*. A higher *removal proportion* corresponds to a lower proportion in the George Floyd dataset, while greater *average transitivity* increases the proportion across both datasets. Higher *semantic similarity to topics* weakens the negative effect of *removal proportion* in the George Floyd dataset and the positive effect of *average transitivity* in the COVID dataset—sometimes even reversing the latter.

Variable	**COVID**	**George Floyd**
Intercept	-0.0151^{*}	-0.0049
	(0.0076)	(0.0053)
Removal Proportion	-0.0088	-0.2124^{***}
	(0.0127)	(0.0451)
u/Automoderator Activity Proportion	-0.0104	0.0073
	(0.0160)	(0.0312)
Standardized Number of Moderators	0.0005	0.0017
	(0.0010)	(0.0012)
Cross-Talk Proportion	0.0055	-0.0057
	(0.0041)	(0.0041)
Average Centralization	-0.0027	0.0029
	(0.0062)	(0.0064)
Average Transitivity	0.0368^{**}	0.0236^{*}
	(0.0113)	(0.0115)
Semantic Similarity to Topics	0.0452	0.0249
	(0.0326)	(0.0357)
Semantic Similarity to Topics * *Removal Proportion*	0.0307	0.1877^{*}
	(0.0794)	(0.0904)
Semantic Similarity to Topics * *u/Automoderator Activity Proportion*	-0.0449	-0.0737
	(0.1048)	(0.0805)
Semantic Similarity to Topics * Standardized *Number of Moderators*	-0.0028	-0.0101
	(0.0057)	(0.0069)
Semantic Similarity to Topics * *Cross-Talk Proportion*	$-0.0494^{\cdot}$	0.0193
	(0.0272)	(0.0280)
Semantic Similarity to Topics * *Average Centralization*	-0.0059	-0.0151
	(0.0400)	(0.0438)
Semantic Similarity to Topics * *Average Transitivity*	-0.2133^{**}	-0.0830
	(0.0772)	(0.0793)
Logged *Volume*	0.0011^{***}	0.0003^{***}
	(0.0001)	(0.00008)
Logged *Size*	-0.0006^{**}	0.2125^{***}
	(0.0002)	(0.0450)
Logged *Average Density*	-0.0024^{**}	0.0034
	(0.0008)	(0.0320)
Content-Based Control Variables (Topics)	Included	Included

Note: $^{\cdot}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

crease from reducing 21.2 messages mentioning unexpected information to reducing 2 such messages every 100 messages. This hypothetical change is significant.

More realistically, within the range of similarity in the George Floyd dataset, Figure 3.2 illustrates that as the similarity increases, the negative effect of *removal proportion* on the *proportion of messages mentioning unexpected information* weakens from around -0.2 to around -0.14, which is not as significant. And it remains negative throughout. To detail how this interaction effect changes the influence of *removal proportion* on the *proportion of messages mentioning unexpected information*, this study uses Figure 3.3 to present five series of predicted *proportion of messages mentioning unexpected information* with variation in *removal proportion*. These series correspond to five different levels of similarity. Figure 3.3 shows similar rates of decrease across all five series. For example, when information's similarity is at its 25th percentile, a community can reduce 7 messages mentioning unexpected information out of 1000 (from 12 to 5) by moving from removing 1 to removing 2 messages per 100. With the similarity at its 75th percentile, the change can induce 8 fewer messages mentioning unexpected information out of 1000 (from 15 to 7). As noted earlier, a median subreddit has around 4000 messages in the four weeks after the murder of George Floyd. Having unexpected information significantly more on-topic is associated with reducing 4 more such messages for a median subreddit, indicating that the influence of the interaction effect of similarity is small.

Overall, the results provided partial support for what this study hypothesized in H1a and H1b. Furthermore, given that the marginal effects of *removal proportion* stay negative as *semantic similarity to topics* increases, the function of governance in regulating sensitivity is in the same direction regardless of whether the unexpected information is on-topic. This trend reflects that online communities' rejection of unexpected information by their formal structures can happen regardless of its topic alignment.

3.2 Cross-talk and Online Communities' Sensitivity to Unexpected Information

H2a focuses on whether cross-talk in online communities is positively associated with those communities' sensitivity to unexpected information. Results in Table 3.1 partially sup-

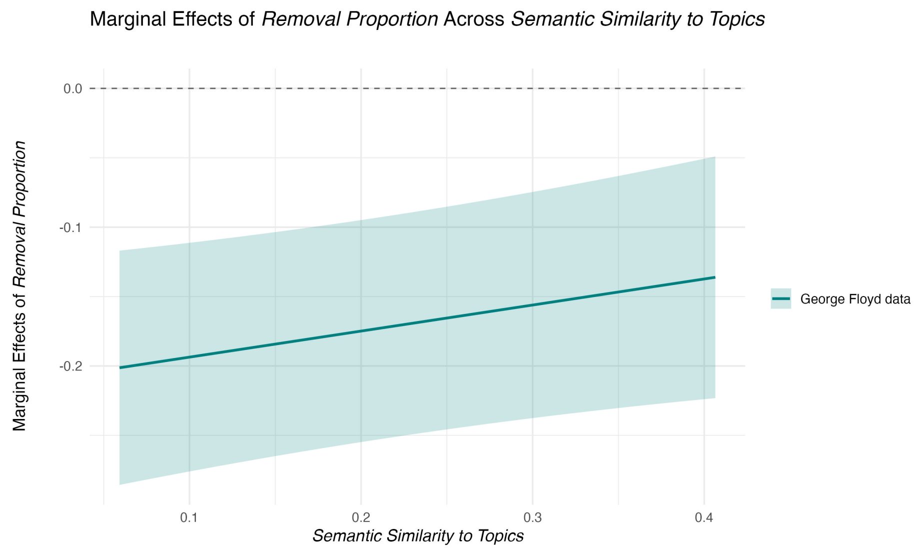

Figure 3.2: Marginal effects of *removal proportion* as a function of unexpected information (COVID)'s *semantic similarity to topics*. This figure presents a positive relationship between the marginal effects of *removal proportion* and the similarity. The marginal effects are always negative in the range of similarity.

port this hypothesis by a significant positive relationship between *average transitivity* and the *proportion of messages mentioning unexpected information* in both datasets ($\beta = 0.0368$, $SE = 0.0113$, $p = 0.0012$ in the COVID dataset, $\beta = 0.0236$, $SE = 0.0115$, $p = 0.0402$ in the George Floyd dataset). Specifically, when the other variables are 0, changing a community from having no connected triangles to having all triads connected as triangles is associated with 3.6 more messages mentioning unexpected information every 100 messages in the COVID dataset and 2.3 more in the George Floyd dataset.

As the illustration of the influence of *average transitivity* on the *proportion of messages mentioning unexpected information* in more realistic situations, Figure 3.4 and Figure 3.5 both present an approximately linear relationship. When the other variables are at their median, every increase in the *average transitivity* by 1%, i.e., one more percent of the triads in interaction networks are connected as triangles across all the posts in a subreddit, corresponds to an average increase in *proportion of messages mentioning unexpected information* of 0.000004 ($\sigma = 0.000003$) in the COVID dataset and 0.0001 ($\sigma = 0.00001$) in the George

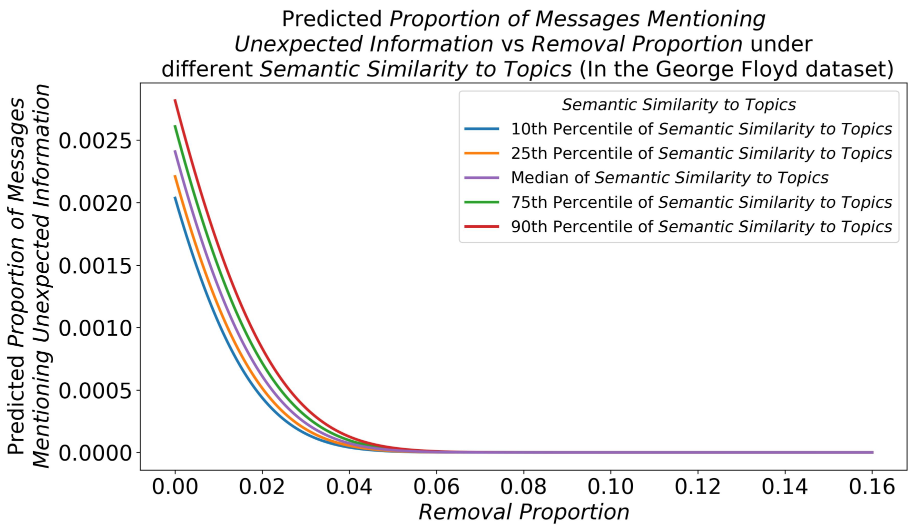

Figure 3.3: Predicted *proportion of messages mentioning unexpected information* as a function of *removal proportion* when the unexpected information (George Floyd)'s *semantic similarity to topics* is at different levels. The relationship between the *proportion of messages mentioning unexpected information* and *removal proportion* is similar across different levels of similarity.

Floyd dataset. Since the median *volume* of messages in the four weeks after the relevant public event is $3,972$ for the COVID dataset and $4,563$ for the George Floyd dataset, one more percent of triads becoming fully connected as triangles can only induce 0.016 more messages mentioning unexpected information for a median subreddit in the COVID dataset while half a message more in the George Floyd dataset. The difference between every 10th percentile of *average transitivity* is roughly 0.01 in both datasets, indicating that encouraging one more percent of triads to reach full connections is a big change for a community. However, such a big change's effect is trivial and demonstrates the low effectiveness of *average transitivity* in influencing the *proportion of messages mentioning unexpected information.*

H2b predicted that unexpected information's topic alignment with a community's conventional topics would moderate the relationship between cross-talk and sensitivity, such that increased topic alignment would weaken the positive effect of cross-talk or strengthen its negative effect. Results in Table 3.1 partially support H2b. Among the three interaction terms between informal structures and *semantic similarity to topics* across the two datasets,

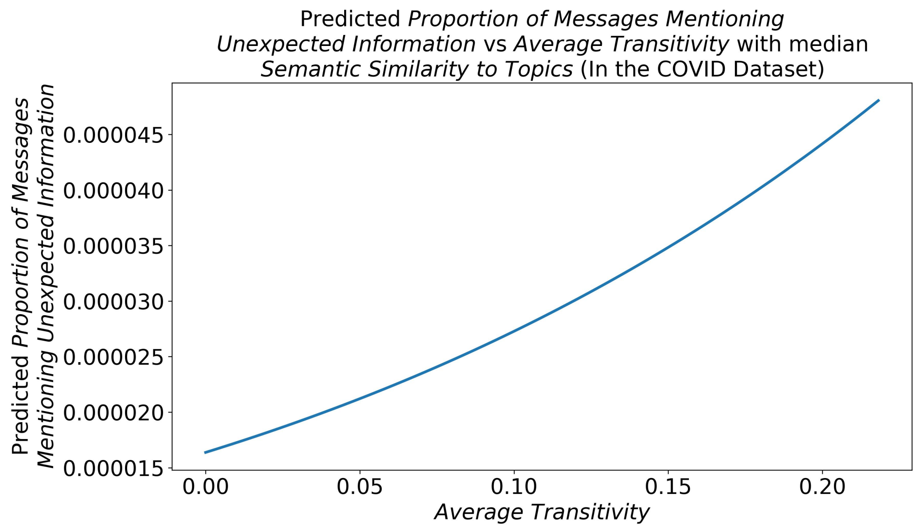

Figure 3.4: Predicted *proportion of messages mentioning unexpected information* as a function of *average transitivity* in the COVID dataset with unexpected information's *semantic similarity to topics* at its median. With median similarity, the *proportion of messages mentioning unexpected information* is positively associated with *average transitivity*. The effect of the association is trivial.

only the interaction between *semantic similarity to topics* and *average transitivity* in the COVID dataset is statistically significant ($\beta = -0.2133,\ SE = 0.0772,\ p = 0.0057$). When the similarity is 0, the marginal effect of *average transitivity* in the COVID dataset is 0.0368, between a community that takes the information completely off-topic and another taking it completely on-topic, the effect of *average transitivity* decrease from increasing 3.6 messages mentioning unexpected information to reducing 17.7 such messages every 100 messages. This hypothetical change is significant.

In more realistic situations, the marginal effect of *average transitivity* decreases from around 0.0368 to -0.0499 from the minimum (0) similarity to its maximum (0.4064) in the COVID dataset. This indicates that with the change from no triads connected to all triads connected, the community that takes COVID as off-topic can have 3.6 more messages out of 100 messages mentioning it, while the community taking it as on-topic has 5 fewer such messages out of 100. As an illustration that specifies the effect of the similarity's interaction effect, Figure 3.7 indicates that the positive association between *average transitivity*

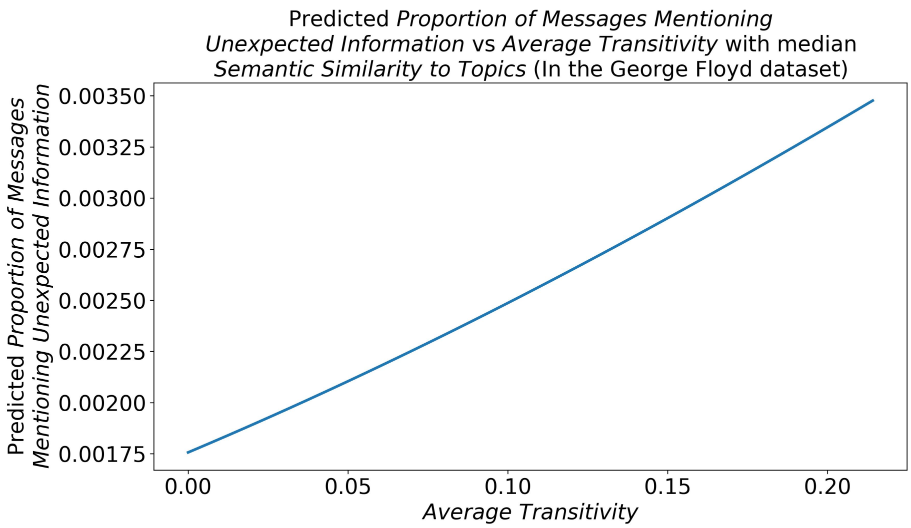

Figure 3.5: Predicted *proportion of messages mentioning unexpected information* as a function of *average transitivity* in the George Floyd dataset with unexpected information's *semantic similarity to topics* at its median. With median similarity, the *proportion of messages mentioning unexpected information* is positively associated with *average transitivity*. The effect of the association is relatively larger than that in the COVID dataset but is still small.

and the *proportion of messages mentioning unexpected information* in the COVID dataset declines in strength as the similarity increases from its 10th percentile (0.1131) to its 75th percentile (0.1641). When the similarity is at its 90th percentile (0.1811), the relationship between *average transitivity* and the proportion becomes negative in Figure 3.7. Therefore, this interaction effect between the similarity and *average transitivity* is considerable. However, because the effect of *average transitivity* on the *proportion of messages mentioning unexpected information* is small, this moderation effect's significance is limited.

Overall, the regression models provide partial support for both H2a and H2b. As one important aspect of cross-talk intensity, *average transitivity* generally promotes the *proportion of messages mentioning unexpected information*. Only when the unexpected information's *semantic similarity to topics* reaches a relatively high level within the COVID dataset does *average transitivity* have a negative relationship with *proportion of messages mentioning unexpected information*. However, the effect of *average transitivity* is relatively small across

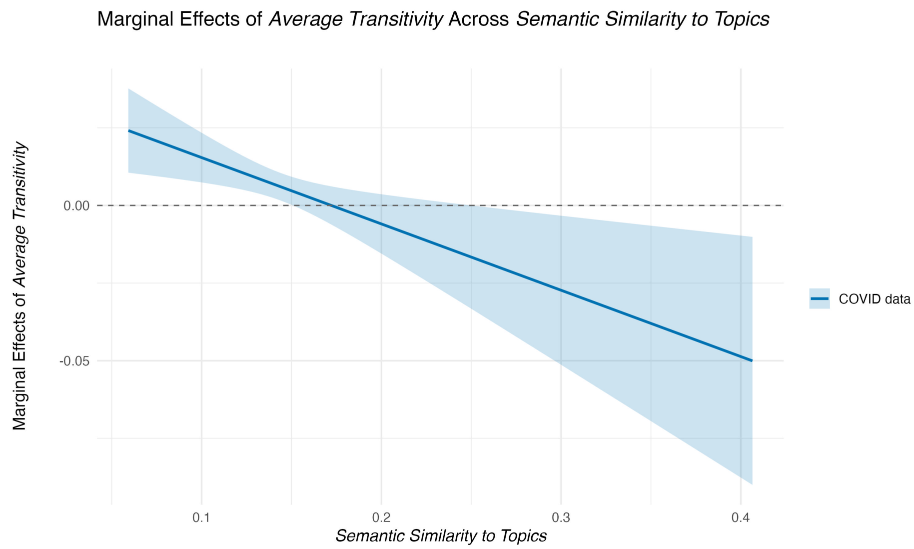

Figure 3.6: Marginal effects of *average transitivity* as a function of unexpected information (COVID)'s *semantic similarity to topics*. This figure presents a negative relationship between the marginal effects of *average transitivity* and *semantic similarity to topics*. The marginal effects are generally positive when the similarity is low but turn negative when the similarity exceeds a certain threshold.

two datasets, which also renders the interaction effects of *semantic similarity to topics* negligible in terms of influencing the *proportion of messages mentioning unexpected information*

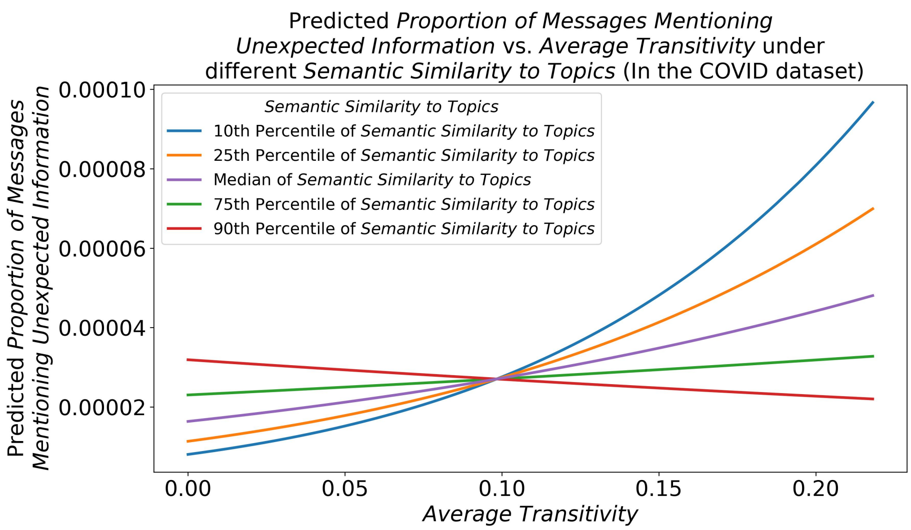

Figure 3.7: Predicted *proportion of messages mentioning unexpected information* as a function of *average transitivity* when the unexpected information (COVID)'s *semantic similarity to topics* is at different levels. The relationship between the *proportion of messages mentioning unexpected information* and *average transitivity* is positive most of the time with strength decreasing as the similarity increases. In this graph, the relationship between *average transitivity* and the *proportion of messages mentioning unexpected information* turns negative when the similarity is at its 90th percentile.

Chapter 4

DISCUSSION

This study investigates variations in online communities' sensitivity to unexpected information from the perspectives of communities' institutions, focusing on formal and informal structures. Using two sets of hypotheses, this study hypothesizes that governance and cross-talk, representing two kinds of community structures, influence sensitivity in opposite directions. The results present partial support from one of the relevant variables for formal and informal structures, respectively, for both sets of hypotheses. Specifically, *removal proportion* in the George Floyd dataset has a negative relationship with the *proportion of messages mentioning unexpected information*, providing support for governance's positive effect on sensitivity. Also, *average transitivity* across two datasets is positively associated with *proportion of messages mentioning unexpected information*, evidencing cross-talk's positive influence on sensitivity. Also, topic alignment's moderating effect finds aligned results. The marginal effect of *removal proportion* increases in terms of absolute value as unexpected information (the murder of George Floyd)'s *semantic similarity to topics* increases, validating the positive interaction effect of topic alignment. On the other hand, an increase in unexpected information (the COVID pandemic)'s *semantic similarity to topics* is associated with the decrease in the marginal effect of *average transitivity*, which supports the negative influence of topic alignment.

In terms of the influence on sensitivity, the effect of governance is considerable, effectively influencing the exposure of unexpected information to community members, and is barely susceptible to the interaction effects by topic alignment. On the contrary, cross-talk's influence on sensitivity is rather trivial, limiting the interaction effect's significance.

In sum, these results suggest that a specific aspect of governance consistently reduces communities' sensitivity to unexpected information with considerable effects, while a particular dimension of cross-talk generally promotes communities' sensitivity to unexpected information with trivial effects.

4.1 Implications

Based on the findings, this study offers three implications for online community research. First, the findings of this research shed light on the reactions of online communities' institutions to one specific kind of off-topic message. Previous studies usually focus on the on-topicness against the disturbance of off-topic messages whose topic alignment is clear to community members (Park et al., 2016). The unexpected information studied here does not present direct off-topicness to online community members. Instead, because of its unprecedentedness, unexpected information can motivate its exclusion in online communities merely by the uncertainty of its on-topicness. Explicating the factors that may influence online communities' sensitivity to unexpected information not only uncovers how community institutions uphold online communities' on-topicness but also connects the findings to online communities' conservativeness when facing uncertainty. According to the results, online communities with stronger governance consistently exhibit lower sensitivity to unexpected information. This pattern suggests that online communities' insistence on on-topicness can take effect merely due to the uncertainty provided by unexpected information's unprecedentedness without clear off-topicness. In contrast, higher levels of cross-talk, which imply lower solidarity with the communities, correspond to a generally stronger tendency to explore uncertain situations, although this positive effect reverses when the unexpected information is highly aligned with the community's conventional topics. Despite the limited size of effects on sensitivity, the way topic alignment influences the correlation between cross-talk and sensitivity reflects the self-organizing discussions' preference for unexpected information is relevant to its off-topicness. Through these analyses, this study unpacks how the emergence of unexpected information interacts with institutions in online communities.

Amid the development of online communities, skepticism about the implications of such digital public spheres also emerges, worrying about the fragmentation that induces echo chambers and polarization (Barberá et al., 2015; Brummette et al., 2018). However, as uncovered by this study, although strong governance that aims to maintain coherent online discussions does inhibit the exploration of unexpected information, informal structures that feature cross-talk can still somewhat contribute to an open environment for unexpected information. With the application of the algorithmic feed by online community platforms, finding and joining interesting conversations in online communities becomes easier, which does not require an understanding of an online community's conventions. Moreover, because such participation tends to be free of the sense of community, it may introduce more cross-talk to online communities. Thus, the influence of formal regulations that may exacerbate echo chambers and polarization weakens. The concerns around online communities' contribution to forming echo chambers may be further diminished.

Second, this study provides evidence for the importance of informal structures in shaping online communities' institutions, which highlights the importance of network structures within online communities. The results provide support for the effect of average transitivity on sensitivity. This distinction reveals that the effects of cross-talk are the strongest when it happens as a group-level behavior. The cross-talk proportion measures the conversational behavior at the dyadic level, only indicating that two persons engage in a conversation that excludes the OP. In comparison, the average transitivity reflects the potential for subgroups to form within an online discussion group. Because triadic closures, as a network structure, provide more stability for subgroup members (Abebe et al., 2022), the formation of these subgroups may reduce members' solidarity with the subreddit by inducing members' dependence on their subgroups. This finding also suggests that micro-level behaviors organized at the group level are effective in shaping the effects of institutions within an online community.

Third, the negative relationship between governance and sensitivity confirms the effectiveness of formal structures within online communities. As this study specifically uncovers the significant effect of *removal proportion*, it illuminates the importance of micro-level behaviors in enabling the effectiveness of formal structures. All three variables reflect the intensity of community governance. However, governance is only becoming more effective when the *removal proportion* increases. Neither a larger number of moderators nor an increase in the proportion of the u/Automoderator activity is associated with a lower sensitivity. This result indicates that deploying more moderators for an online community does not necessarily contribute to fostering formal structures that maintain on-topicness within the community. Besides, as pointed out in previous research (Kuo et al., 2023), moderation is a complicated task. Relying on scripts that can only process simple tasks, it is difficult for the u/Automoderator to perform well on moderation all the time. The results in this study show that the increase in the number of tasks finished by the u/Automoderator is not sufficient in constructing effective formal structures within online communities, highlighting the importance of contributions from human moderators.

Thus, what makes formal structures effective is the micro-level behaviors of human moderators who are responsible for practicing governance. Generally, the implementation of online communities' governance uses rules written by community administrators, and most rules contain the maintenance of on-topicness (Reddy and Chandrasekharan, 2023). The moderators are expected to implement the rules in their work. However, not all moderators have the time and energy to consistently contribute to the implementation of community rules. Various factors limit the capability of moderators to ensure the effectiveness of community governance. For example, the lack of standardized training for moderators may cause them to understand the requirements for their moderation work differently (Seering et al., 2019). Eventually, what contributes to the construction of an institution is the visible behaviors of moderators. Community members only present behavioral variations when the amount of governance practiced in each community varies. The formal structures, despite often being artifacts like written rules, require informal behavioral patterns to induce influence.

Also, the results in this study validate findings in previous research that community members learn about the norms of online communities by observing and engaging with activities within online communities (Gallagher and Savage, 2015). The indispensability of micro-level behaviors in constructing the institutions within online communities calls for a question about the essence of online communities. The concept of online communities describes them as online spaces featuring boundaries and membership, though the clarity of these artifacts may vary (Plant, 2004). Those artifacts are typically rules and community descriptions that are written on associated webpages. However, artifacts can only accomplish their expected influence when the micro-level behaviors match the intentions behind them. It seems to suggest that online communities are the manifestation of their members' behaviors rather than artifacts defined by the intention of their founders or administrators. As algorithmically driven content feeds increasingly direct people to discussions of interest in communities whose norms these people are not familiar with, the essence of online communities may become more ambiguous, making it more difficult to normalize members' behaviors. Building an online community with desired features thus becomes a more complicated job, which invites further research into the understanding of online communities under new circumstances.

4.2 Threats to Validity

This study implements censored regression to address the large proportion of zero values within the datasets this study analyzes. However, due to the large volume of zeros, this study has not yet found a way to meet the normality assumption for linear regression, as the residuals tend to be skewed by the influence of the large volume of zero values. The results also present a nonlinear relationship between the independent variables and the predicted values of sensitivity. To overcome this threat, the next steps of this study plan are to explore using the number of comments, including corresponding keywords, to construct regression models with count data or other models that feature different assumptions.

Besides, when this study implements the BERT-based sentence transformers to compute the similarity between the phrases that describe relevant public events and online communities' conventional topics, the performance of this method is not validated. Although trained using up-to-date content that includes online discussion around the COVID-19 pandemic and the murder of George Floyd, the all-mpnet-base-v2 model's performance on this specific task is unknown. Next steps to overcome this threat to validity will focus on manually examining the model outputs or hiring human coders to rate the accuracy of the current approach.

4.3 Future Directions

Currently, this study only analyzes online communities' reactions to unexpected information associated with the COVID-19 pandemic and the murder of George Floyd. Compared with George Floyd's murder, which incited more reactions in the U.S. than in other areas worldwide, COVID-19 is a much more impactful event and can trigger comparably significant discussions globally. As subreddits do not only focus on events happening in the United States, events more globally known than the murder of George Floyd may better reflect how online communities worldwide react to unexpected information. Future studies may select more comparable events to study the association between online communities' institutions and sensitivity to unexpected information.

Also, this research only looks into how the sensitivity of online communities with top activeness is associated with communities' formal and informal structures. However, when the activeness of online communities decreases, community members' actions may have different implications on the influence of formal and informal structures. For example, less active communities may be populated by inactive posts and are less amenable to the influence of interaction patterns like cross-talk. Analyzing how micro-level behaviors can provide more depth into the processes where online communities construct their institutions. This study would suggest further research look into online communities with distinct levels of activeness to find out how online communities that are active to different degrees react to unexpected information that may disrupt communities' on-topicness.

Lastly, this study does not emphasize a significant distinction between online communities, which is the distinction between the generalist and specialist online communities (Waller and Anderson, 2019). Though this study controlled potential effects of this distinction by implementing content-based control variables, this distinction may have further implications for online communities' institutions. With the involvement of algorithmic feeds, generalist online communities may contain more activities from people who are not the communities' members guided by the feed. This tendency blurs the boundaries of online communities to different degrees, inducing varying influences from communities' formal and informal structures. Future research exploring such a distinction can advance the knowledge of online communities' institutions with the development of information and communication technologies.

4.4 Conclusion

Online communities are idealized venues for people with shared interests or goals to engage in discussions around topics of interest supported by the communities' institutions. However, the effectiveness of such institutions varies across communities, inducing communities' varying sensitivity to unexpected information. By leveraging large-scale data analysis, this study investigated how formal and informal structures of online communities influence communities' sensitivity to unexpected information, with the interaction effects of the information. Results provide partial support for the tendency that increased governance inhibits online communities' sensitivity to unexpected information with considerable effect, while more cross-talk generally promotes sensitivity with limited effects. An increase in cross-talk only reduces sensitivity when unexpected information is highly aligned with a community's conventional topics. Such findings suggest the high effectiveness of online communities' formal structures and the noticeable role informal structures play in shaping the effectiveness of online communities' institutions. Furthermore, this study's results illuminate the importance of micro-level behaviors in constructing institutions, inviting further investigations into how informal and formal structures implicate members' behaviors in an online community.

BIBLIOGRAPHY

Abebe, R., Immorlica, N., Kleinberg, J., Lucier, B., and Shirali, A. (2022). On the Effect of Triadic Closure on Network Segregation. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 249–284, Boulder CO USA. ACM.

Adler, P. S. and Borys, B. (1996). Two Types of Bureaucracy: Enabling and Coercive. *Administrative Science Quarterly*, 41(1):61.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., and Bonneau, R. (2015). Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, 26(10):1531–1542. Publisher: SAGE Publications Inc.

Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., and Blackburn, J. (2020). The Pushshift Reddit Dataset. *Proceedings of the International AAAI Conference on Web and Social Media*, 14:830–839.

Benkler, Y., Roberts, H., Faris, R., Solow-Niederman, A., and Etling, B. (2015). Social Mobilization and the Networked Public Sphere: Mapping the SOPA-PIPA Debate. *Political Communication*, 32(4):594–624.

Berger, P. L. and Luckmann, T. (1966). *The Social Construction of Reality*. Doubleday & Company, New York.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022.

Brummette, J., DiStaso, M., Vafeiadis, M., and Messner, M. (2018). Read All About It: The Politicization of "Fake News" on Twitter. *Journalism & Mass Communication Quarterly*, 95(2):497–517. Publisher: SAGE Publications Inc.

Burgoon, J. K. (2015). Expectancy Violations Theory. In *The International Encyclopedia of Interpersonal Communication*, pages 1–9. John Wiley & Sons, Ltd. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118540190.wbeic102.

Burgoon, J. K. and Jones, S. B. (1976). Toward a Theory of Personal Space Expectations and Their Violations. *Human Communication Research*, 2(2):131–146.

Chaiken, S. and Ledgerwood, A. (2012). A theory of heuristic and systematic information processing. In *Handbook of theories of social psychology*, pages 246–266. Sage Publications Ltd, Thousand Oaks, CA.

Coleman, J. S. (1990). *Foundations of social theory.* Belknap Press of Harvard University Press, Cambridge, Mass.

Dekker, D., Krackhardt, D., and Snijders, T. A. B. (2019). Transitivity correlation: A descriptive measure of network transitivity. *Network Science*, 7(3):353–375.

DiMaggio, P. J. and Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48(2):147–160.

Duguay, P. A. (2022). Read it on Reddit: Homogeneity and Ideological Segregation in the Age of Social News. *Social Science Computer Review*, 40(5):1186–1202. Publisher: SAGE Publications Inc.

Ferree, M. M. and Gamson, W. A. (2002). Four Models of the Public Sphere in Modern Democracies. *Theory and Society*, 31(3):289–324.

Fiesler, C., Jiang, J., McCann, J., Frye, K., and Brubaker, J. (2018). Reddit Rules! Characterizing an Ecosystem of Governance. *Proceedings of the International AAAI Conference on Web and Social Media*, 12(1):72–81. Number: 1.

Fujimoto, M. (2016). Team Roles and Hierarchic System in Group Discussion. *Group Decision and Negotiation*, 25(3):585–608.

Gallagher, S. E. and Savage, T. (2015). "What is, Becomes What is Right": A Conceptual Framework of Newcomer Legitimacy for Online Discussion Communities. *Journal of Computer-Mediated Communication*, 20(4):400–416.

Gilson, L. L., Maynard, M. T., Jones Young, N. C., Vartiainen, M., and Hakonen, M. (2015). Virtual Teams Research: 10 Years, 10 Themes, and 10 Opportunities. *Journal of Management*, 41(5):1313–1337.

Granovetter, M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology*, 91(3):481–510.

Grimmelmann, J. (2015). The Virtues of Moderation. *Yale Journal of Law and Technology*, 17(42). Accepted: 2021-11-26T11:57:08Z.

Habermas, J. (1962). The Structural Transformation of the Public Sphere.

Hechter, M. (1987). *Principles of Group Solidarity.* University of California Press.

Henningsen, A. (2010). Estimating Censored Regression Models in R using the censReg Package. *R Package Vignettes*, 5(12):1–12.

Hwang, S. and Foote, J. D. (2021). Why do People Participate in Small Online Communities? *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–25.

Jarman, R. (2005). When Success Isn't Everything – Case Studies of Two Virtual Teams. *Group Decision and Negotiation*, 14(4):333–354.

Kiesler, S., Kraut, R. E., Resnick, P., and Kittuer, A. (2012). Regulating behavior in online communities. In *Building Successful Online Communities: Evidence-Based Social Design*, pages 125–177. MIT Press. Google-Books-ID: lIvBMYVxWJYC.

Kollock, P. and Smith, M. (1996). Managing the virtual commons: Cooperation and conflict in computer communities. In Herring, S. C., editor, *Pragmatics & Beyond New Series*, volume 39, pages 109–128. John Benjamins Publishing Company, Amsterdam.

Kuo, T., Hernani, A., and Grossklags, J. (2023). The Unsung Heroes of Facebook Groups Moderation: A Case Study of Moderation Practices and Tools. *Proc. ACM Hum.-Comput. Interact.*, 7(CSCW1):97:1–97:38.

Lawler, E. J., Thye, S. R., and Yoon, J. (2009). *Social Commitments in a Depersonalized World*. Russell Sage Foundation.

Lin, Z., Salehi, N., Yao, B., Chen, Y., and Bernstein, M. (2017). Better When It Was Smaller? Community Content and Behavior After Massive Growth. *Proceedings of the International AAAI Conference on Web and Social Media*, 11(1):132–141. Number: 1.

Malinen, S. (2015). Understanding user participation in online communities: A systematic literature review of empirical studies. *Computers in Human Behavior*, 46:228–238.

Meyer, J. W. and Rowan, B. (1977). Institutionalized Organizations: Formal Structure as Myth and Ceremony. *American Journal of Sociology*, 83(2):340–363.

Newman, M. E. J., Strogatz, S. H., and Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(2):026118.

North, D. C. (1991). Institutions. *Journal of Economic Perspectives*, 5(1):97–112.

Op 'T Roodt, H., Krug, H., and Otto, K. (2021). Subgroup Formation in Diverse Virtual Teams: The Moderating Role of Identity Leadership. *Frontiers in Psychology*, 12:722650.

Panzarasa, P., Opsahl, T., and Carley, K. M. (2009). Patterns and dynamics of users' behavior and interaction: Network analysis of an online community. *Journal of the American Society for Information Science and Technology*, 60(5):911–932. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.21015.

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2):259–283. Place: US Publisher: Sage Publications.

Park, A., Hartzler, A. L., Huh, J., Hsieh, G., McDonald, D. W., and Pratt, W. (2016). "How Did We Get Here?": Topic Drift in Online Health Discussions. *Journal of Medical Internet Research*, 18(11):e284.

Plant, R. (2004). Online communities. *Technology in Society*, 26(1):51–65.

Prell, C., Reed, M., Racin, L., and Hubacek, K. (2010). Competing Structure, Competing Views: The Role of Formal and Informal Social Structures in Shaping Stakeholder Perceptions. *Ecology and Society*, 15(4):art34.

Reddy, H. and Chandrasekharan, E. (2023). Evolution of Rules in Reddit Communities. In *Computer Supported Cooperative Work and Social Computing*, pages 278–282, Minneapolis MN USA. ACM.

Ren, Y. and Kraut, R. E. (2014). Agent-Based Modeling to Inform Online Community Design: Impact of Topical Breadth, Message Volume, and Discussion Moderation on Member Commitment and Contribution. *Human–Computer Interaction*, 29(4):351–389. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07370024.2013.828565.

Roethlisberger, F. J. and Dickson, W. J. (1939). *Management and the worker*. Management and the worker. Harvard Univ. Press, Oxford, England. Pages: xxiv, 615.

Röder, M., Both, A., and Hinneburg, A. (2015). Exploring the Space of Topic Coherence Measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, WSDM '15, pages 399–408, New York, NY, USA. Association for Computing Machinery.

Salton, G. and Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5):513–523.

Schwämmlein, E. and Wodzicki, K. (2012). What to Tell About Me? Self-Presentation in Online Communities. *Journal of Computer-Mediated Communication*, 17(4):387–407.

Seering, J., Wang, T., Yoon, J., and Kaufman, G. (2019). Moderator engagement and community development in the age of algorithms. *New Media & Society*, 21(7):1417–1443.

Shrum, W. and Mullins, N. (1988). CHAPTER 4 - NETWORK ANALYSIS IN THE STUDY OF SCIENCE AND TECHNOLOGY. In Van raan, A. F. J., editor, *Handbook of Quantitative Studies of Science and Technology*, pages 107–133. Elsevier, Amsterdam.

Siino, M. (2024). All-Mpnet at SemEval-2024 Task 1: Application of Mpnet for Evaluating Semantic Textual Relatedness. In Ojha, A. K., Doğruöz, A. S., Tayyar Madabushi, H., Da San Martino, G., Rosenthal, S., and Rosá, A., editors, *Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024)*, pages 379–384, Mexico City, Mexico. Association for Computational Linguistics.

Smith-Doerr, L. and Powell, W. W. (2005). Networks and Economic Life. In Smelser, N. J. and Swedberg, R., editors, *The Handbook of Economic Sociology, Second Edition*, pages 379–402. Princeton University Press, stu - student edition edition.

Srinivasan, K. B., Danescu-Niculescu-Mizil, C., Lee, L., and Tan, C. (2019). Content Removal as a Moderation Strategy: Compliance and Other Outcomes in the ChangeMyView Community. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–21.

Tausczik, Y. and Huang, X. (2019). The Impact of Group Size on the Discovery of Hidden Profiles in Online Discussion Groups. *Trans. Soc. Comput.*, 2(3):10:1–10:25.

Waller, I. and Anderson, A. (2019). Generalists and Specialists: Using Community Embeddings to Quantify Activity Diversity in Online Platforms. In *The World Wide Web Conference*, WWW '19, pages 1954–1964, New York, NY, USA. Association for Computing Machinery.

Weld, G., Leibmann, L., Zhang, A. X., and Althoff, T. (2025). Perceptions of Moderators as a Large-Scale Measure of Online Community Governance. arXiv:2401.16610 [cs].

Yu, Y., Jiang, J., and Dhillon, P. S. (2024). Characterizing the Structure of Online Conversations Across Reddit. *Proc. ACM Hum.-Comput. Interact.*, 8(CSCW2):374:1–374:23.