

Reporting the Community Beat: Practices for Moderating Online Discussion at a News Website

BRIAN MCINNIS, University of California, San Diego, USA

LEAH AJMANI, Cornell University, USA

LU SUN, University of California, San Diego, USA

YIWEN HOU, University of California, San Diego, USA

ZIWEN ZENG, University of California, San Diego, USA

STEVEN P. DOW, University of California San Diego, USA

Due to challenges around low-quality comments and misinformation, many news outlets have opted to turn off commenting features on their websites. The New York Times (NYT), on the other hand, has continued to scale up its online discussion resources to reach large audiences. Through interviews with the NYT moderation team, we present examples of how moderators manage the first ~24 hours of online discussion after a story breaks, while balancing concerns about journalistic credibility. We discuss how managing comments at the NYT is not merely a matter of content regulation, but can involve reporting from the “community beat” to recognize emerging topics and synthesize the multiple perspectives in a discussion to promote community. We discuss how other news organizations—including those lacking moderation resources—might appropriate the strategies and decisions offered by the NYT. Future research should investigate strategies to share and update the information generated about topics in the news through the course of content moderation.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**.

Additional Key Words and Phrases: content moderation, news outlet, community reporting

ACM Reference Format:

Brian McInnis, Leah Ajmani, Lu Sun, Yiwen Hou, Ziwen Zeng, and Steven P. Dow. 2021. Reporting the Community Beat: Practices for Moderating Online Discussion at a News Website. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 333 (October 2021), 25 pages. <https://doi.org/10.1145/3476074>

1 INTRODUCTION

A central challenge for any online community is how to moderate the content that users generate and share with others in the community [49]. Grimmelmann [40] defines online community moderation as a series of “governance mechanisms that structure participation to facilitate cooperation and prevent abuse” [40, pg. 42]. In practice, moderating an online community may involve large teams of people working in conjunction with automated systems to keep up with the ebb and flow of user-generated content [33]. Research about content moderation has continued to develop from notions of how to organize content (e.g., deleting, filtering, recommending, synthesizing) [40], to strategies about how to advance broad social and content norms that address insularity [2],

Authors’ addresses: Brian McInnis, bmcinnis@ucsd.edu, University of California, San Diego, USA; Leah Ajmani, lha37@cornell.edu, Cornell University, USA; Lu Sun, l5sun@ucsd.edu, University of California, San Diego, USA; Yiwen Hou, yih267@ucsd.edu, University of California, San Diego, USA; Ziwen Zeng, ziz236@ucsd.edu, University of California, San Diego, USA; Steven P. Dow, spdown@ucsd.edu, University of California San Diego, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2573-0142/2021/10-ART333 \$15.00

<https://doi.org/10.1145/3476074>

offer care [21, 57, 87], champion free speech [48], and encourage governance structures that are community-centered [69]. Content moderation is meant to promote productivity and openness within an online community [18, 40], but in the context of a news website, moderators may also need to pay careful attention to how contributions affect the reporting [3, 12, 74] and reporters' credibility [20, 80].

Many news outlets have tried to foster online discussion around news articles (e.g., NPR [44], Popular Science [50], Reuters [23], Vice News [76]), but have decided to shutdown their commenting sections. Several factors played into this. First, organizations incur substantial financial and human costs associated with moderation [23, 34, 68]. Second, people who actively participate tend to reflect a small segment of the news audience [44]. Third, online comments about news articles often exhibit low quality and toxic elements (e.g., including profanity, personal attacks) [7, 17, 30]. Placing toxic comments alongside a news article can also distort what a person remembers about the article, which can polarize their views of the issue [3] and detract from the reporting [20]. For many news outlets, the costs of content moderation outweigh the benefit of hosting a discussion [23, 44]. Additionally, the news industry as a whole has experienced a steady decline in revenue following the 2007-2009 global recession, which has resulted in substantial cuts to newsroom staff and infrastructure [5, 39].

The New York Times (NYT), on the other hand, has continued to reach new audiences by experimenting with new ways to report and tell stories online. The NYT website went live on January 22, 1996 [55]. Over time the number of articles “open” at the NYT website for ~24 hours of online discussion has gradually scaled up—from one article per-day in 2006 [11], to 8-10 articles [10], and then 20 [24]. In 2017, commenting was enabled on almost all articles, blogs, and learning initiatives hosted at the website [25]. Enhancing the online community infrastructure has been part of a broad initiative to make paid subscriptions to online content a central component of the business model [83]. Since 2012, the number of digital-only subscriptions to the NYT has increased from ~566,000 [81] to 6.69 million subscribers, which reflects a large share of the newspaper's annual revenue (\$167 million or ~32% of total revenue as of 2020) [82]. While the digital experience of news has become an increasingly important component of the NYT business model, less is known about how moderators at the NYT have navigated the key challenges facing online discussion that have led other news outlets to turn off commenting features on their websites.

Research Question: What moderation practices has the New York Times created to balance concerns about credibility in reporting, while facilitating a large volume of online discussion within the fast-paced news cycle?

To explore this question, we conducted interviews with seven people who moderate discussion at the NYT website. The findings present examples of the information work that moderators perform by themselves, as a team, and with the audience, in order to promote high-quality online discussions. While the crux of moderating online discussion is about deciding what content to share with a community, we illustrate how NYT moderators apply journalistic techniques to keep pace with the news cycle. These techniques include recognizing emerging topics, investigating dubious claims, coordinating with others across the news organization, and synthesizing the multiple perspectives in a discussion. By reporting from the *community beat*, the moderation team has enabled the NYT to update articles based on personal experiences shared in comments [32], to connect journalists with leads that surface from the discussion [47], and to honor former community members [84]. By participating with the discussion in such a central, yet hidden way, the NYT moderators develop a familiarity with specific people and personalities within the discussion, helping them to recognize community at the news website.

Our findings illustrate a set of community-oriented moderation practices at a news website with ample resources. We discuss how it may not be feasible for moderation teams at other news outlets to adopt the NYT practices. Volunteer community moderators, who regularly facilitate online discussions about news articles at platforms like Reddit, may also be hard pressed to implement all of the NYT practices. We consider how these practices might generalize to other organizations, and how we might disseminate moderation decisions across a discussion ecosystem. Building on existing ideas about cross-community moderation [16], such a model could enable news websites and online communities with relatively few moderation resources to benefit. Drawing examples from our findings, we suggest ways to coordinate moderation resources around news, such as popular topics, valued perspectives, and misinformation that emerged in earlier discussions about an article. The process of coordinating moderation resources across multiple communities might also provide news outlets with a way to monitor how discussions develop in social media and in online forums beyond the publisher's website.

2 LITERATURE REVIEW

2.1 Moderators advance social norms by making decisions on content

There are many reasons to moderate user-generated content for an online community [49]. Content moderation can help to promote productivity and openness, while unmoderated discussion can lead to content that is redundant or that violates community standards [40]. It is not uncommon for large online communities, like those at Reddit, to utilize rules [28], people [14], and technology in the process of moderation (e.g., *Automoderator* [45], *CommentIQ* [64]) and policy-making (e.g., *CivilServant* [59]).

Grimmelmann [40] defines content “organizing” as a collection of moderation tactics used to limit low-quality user-generated content and promote high-quality contributions [40, pp. 58-61]. Simple content organizing includes deleting, editing, annotating, filtering, formatting, and synthesizing. For example, moderators can remove low-quality contributions before they are shared with a community. Removing content may also promote high-quality contributions by reducing the incentive for users to generate content that is unlikely to be shared.

Every moderation decision has consequences, but the decisions and their consequences may be more or less visible to users. For example, deviant communities will try to thwart rule-based moderation systems by varying the language they use to organize followers and dodge enforcement [15]. While increasing the visibility of the rules may enable deviance, presenting the rationale related to specific moderation decisions can promote understanding about community policies. Jhaver et al. [46] argue that providing direct feedback to offending users can promote positive social behaviors within a community.

There are also many approaches to moderation. Grimmelmann [40] describes several dimensions of the moderation system design space: i.e., manually versus automatically, transparently versus secretly, ex post versus ex ante, central versus distributed. For example, the simple task of determining whether a contribution should be rejected could be performed by a human reviewing each contribution manually or by a system applying a predefined list of prohibited terms automatically. Recent advancements have helped to broaden the scope of content moderation from a platform-centered and regulatory perspective [40] to a perspective that centers on community-driven decision-making [69] within an increasingly complex political landscape [37, 38].

2.2 Content moderation decisions are made by people, teams, and technology

Grimmelmann [40] describes how content moderation software can enforce rule-based decisions faster than a human moderator. Automated rule-based systems, like *Automoderator*, can apply

community specific policies (e.g., profanity, malformed post titles, malicious links) to flag content that potentially violates community standards for human moderators to review [45]. Human moderators also use software to search for high-quality comments. As an example, the *CommentIQ* system presents human moderators with a flexible dashboard enabling them to filter, aggregate, and visualize multiple attributes related to a collection of news website comments [64].

Moderation processes may also involve numerous organizational structures. For example, moderators might coordinate as a distributed group, using scoring systems to weigh the independent judgements of several moderators [51, 52]. Some subtle violations of community norms may require group discussion [26]. A community might organize moderators into various configurations (e.g., ad hoc juries [26], teams [22, 70]). Additionally, moderators are situated within the social organization of their community [69] and platform [34, 68].

In commercial platforms, content moderation can involve globally distributed teams of mostly freelance workers who have minimal influence over moderation policy-making [34]. Often moderators on commercial platforms are hidden listeners who—by tuning in and out of conversations across multiple layers of infrastructures, actors, channels, and time—can heavily influence a users’ environment without being recognized [13]. Despite their limited agency at some commercial platforms, Roberts [68] argues that due to the nature of their work with potentially viral content, moderators have some power to permit offensive, yet pivotal content to reach the community.

Within an online community the moderators are often viewed as “leaders” [22, 29, 58, 70, 86]. Unlike the secrecy surrounding moderation policy-making at a commercial platform, many online communities will openly discuss moderation. *CivilServant* [59] is a system that online communities can use to facilitate deliberation about moderation policies and system design by instituting formal testing, debrief, and reporting procedures.

In addition to community moderators, journalists may also be active within an online discussion, particularly around newsworthy events [40]. For example, after Osama Bin Laden’s death, journalists took the lead in disseminating information on Twitter [43]. While not formal moderators, journalists can play an important role in preventing misinformation during a crisis event by responding early to rumors posted on social media [4].

Given the pacing and journalistic concerns of news websites, our paper investigates the organizational structures used to coordinate moderation practices around news-related discussions: What organizational structures are in place to coordinate discussion moderation at the NYT website?

2.3 Moderating an online discussion can be time intensive

There are many considerations when developing a content moderation strategy, but most relevant to a news website is how the volume and speed of the discussion influence moderation. Grimmelmann [40] suggests that moderation can either occur before user-generated content is shared with the community (called “ex ante”) or after the content has been shared (“ex post”). While ex ante moderation has the benefit of removing potentially offensive content to prevent harm, it imposes a delay on sharing content generally. Ex post moderation does not include such a delay, but policing a large discussion for offending content can take time. It may also take time for a community to heal after a specific harm has been felt by its members.

The volume of user-generated content fluctuates over time. Some discussion topics can elicit a burst in user-generated contributions, which can suddenly increase the demand for moderation. Leavitt and Clark [53] describe event-driven communities that form around breaking news, such as the *r/sandy* subreddit. As the hurricane pummelled the eastern seaboard, community moderators at *r/sandy* coordinated their efforts to handle a flood of user-generated content. The *r/sandy* community remained active long after Hurricane Sandy had passed, in part because the volume of user-generated content had subsided to a manageable stream [53].

Online discussion at news websites may offer a unique example of event-driven moderation as the relevant topics continue to develop throughout the fast-paced news cycle. To explore this, our study investigates: What activities do moderators at the NYT routinely perform to facilitate discussion around fast-paced events?

2.4 Journalistic standards play into moderation practices at news websites

The reputation of a news outlet depends on the credibility of its reporting. Online discussion can offer a valuable opportunity for people to engage with the multiple perspectives surrounding a story, but commenting sections can also provide space for people to undermine the reporting credibility [20, 31]. News outlets go through a process of *gatekeeping* news as a way to sustain their credibility [72, 73, 85]. Information about an event must pass through several filters (or “gates”), where reporters, editors, and other “gatekeepers” make choices about what information to include in publication [73]. While some stories, like breaking and timely news, might move quickly through the process, the recursive cycle of gatekeeping in journalism is designed to diminish the influence of any individual gatekeeper on what gets published [72].

There is substantial debate about whether a high standard of discourse is actually possible in commenting sections associated with online news articles [12, 19, 20, 41, 74, 75]. Presenting comments alongside an article can detract from the rigorous gatekeeping involved with the reporting [12, 74], particularly if the comments are toxic [3]. Sometimes journalists get involved with an online discussion, which can lead to more civil and relevant contributions from the community [80]; however, doing so runs the risk of journalists getting coaxed into an open debate about their personal views on an issue, which may impact their reputation [20, 74].

Gatekeeping processes are also central to how a news outlet upholds its legal responsibilities related to publishing [8], e.g., personal privacy, libel, obscenity, inciting violence. However, the commenting sections at news websites are situated within a legal grey area where content moderation is regulated the same way as social media platforms [9, 35]. In the United States of America, the Communications Decency Act (CDA) (1996) Section 230 provides a “safe harbor” for interactive web services. Service providers are not liable for user-generated content in the same way that news outlets are accountable for published statements. Additionally, the act of moderating user-generated content does not make the service into a “publisher” of the content, like a newspaper [9, 35].

User-generated content moderation is meant to promote productivity and openness within an online community [18, 40], but in the context of a news website, moderators may also need to pay careful attention to how contributions engage with the reporting [3, 12, 74] and affect reporters’ credibility [20, 80]. In this paper, we explore how moderation practices are intertwined with publishing standards and journalistic concerns at a major news website.

3 BACKGROUND ON MODERATION AT THE NEW YORK TIMES WEBSITE

This paper presents results from a series of interviews with moderators at the NYT about the steps they take before, during, and after an online discussion as they select and elevate user-generated comments about news articles, blogs, and other initiatives. Figure 1 is a representation of the commenting interface as of April 2021, noting several major components. Only registered users are able to post, reply to, recommend, and flag comments, but any reader can read the comments and generate a hyperlink to *share* a specific comment. As new comments are added, the commenting interface will change over time to include tabs for comments recommended by readers (i.e., Reader Picks), as well as comments recommended by moderators (i.e., NYT Picks).

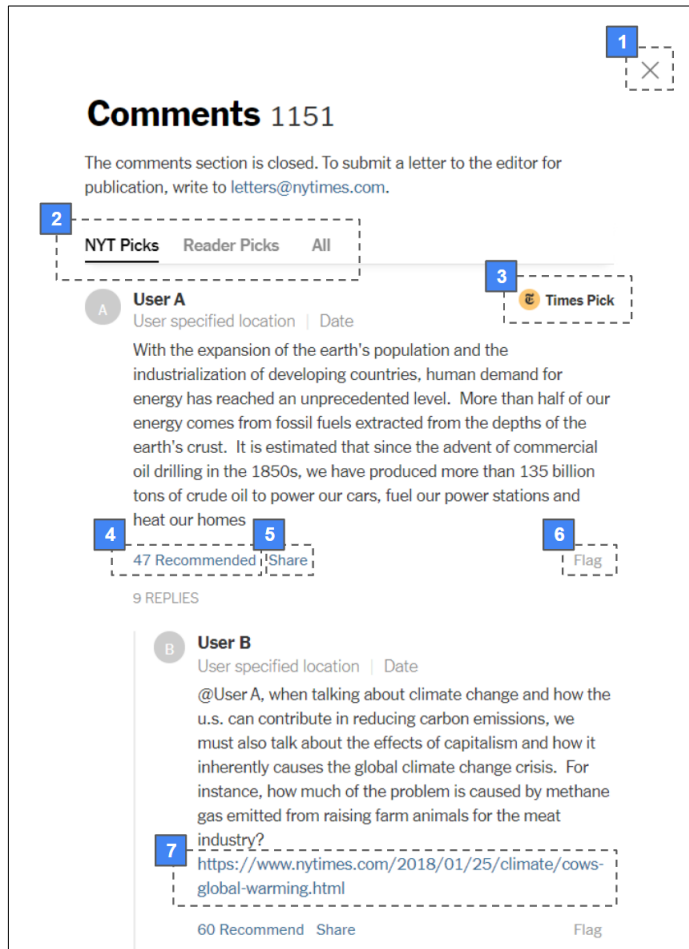


Fig. 1. Representation of the NYT Commenting Interface, including the following features: (1) Users can open and close the interface, which appears along the right side of an article, (2) comments are listed under two standard tabs (i.e., Reader Picks, All), (3) when the NYT Community Team has recommended a comment, the comment receives the *Times Pick* badge and appears under the *NYT Picks* tab, (4) user recommended comments are listed in the *Reader Picks* tab, sorted from most to least recommended, (5) users can also *share* discussion posts with a NYT generated permanent hyperlink for each comment, (6) when users click the comment *flag* button, they are asked to fill out a short checkbox form to indicate the inappropriate elements of the comment: i.e., inflammatory, off topic, personal attack, vulgar, spam, (7) user-generated comments often include text as well as hyperlinks. The comments presented for User A and User B were generated by the research team for illustrative purposes.

The NYT guidelines for commenting are modeled after the NYT standards for publishing news.¹ In a 2008 response to readers, Heather Moore, the Community Editor at the time, described the paper's aspiration to be “the anti-Jerry Springer” of online discussion [61]:

“We follow the same guiding principles that The Times has always followed in print. From our Manual of Style and Usage: *The Times differentiates itself by taking a stand for*

¹NYT Guidelines for Commenting (07 Sep 2020): <https://help.nytimes.com/hc/en-us/articles/115014792387-Comments>

civility in public discourse. While we don't edit readers' words, we do carefully mold their conversations by rejecting personal attacks, incoherence and vulgarity." [61]

In practice, the NYT has applied this standard by developing an *ex ante* moderation process where all comments are reviewed by members of the NYT Community Team (primarily) by hand before publishing them to the news website [11, 66]. Moderators promote these standards by filtering out personal attacks and obscenity, while filtering for comments that are thoughtful, diverse, and succinct. However, the topic of an article matters, as explained by former Community Editor, Bassey Etim: "Any of these guidelines can be bent or even broken, if the article calls for it. And that's where I come in—to decide when to break the rules" [24].

During the past several years, all online discussions at the NYT website have been managed by a team of ~10-15 moderators, with at least a few moderators on at all time to tend to any open discussions [24, 25]. Most articles are only "open" for online discussion during a ~24 hour period. At least since 2010, the largely manual moderation process has also incorporated automated text classification systems that take a first-pass at flagging abusive comments [10]. In 2017, the NYT incorporated the Google *Moderator AI* system, which enables the moderation team to filter for comments based on a predicted "toxicity score" [25].

Several of the blogs and learning initiatives have developed additional moderation steps to facilitate discussion for their communities. For example, since 2017 the NYT Learning Network [63] has partnered with the American Statistical Association (ASA) to facilitate a weekly online discussion with primary and secondary students about articles that contain data visualizations [36]. While moderators are not typically visible in online discussion at other parts of the NYT website, some play an active role in the weekly "What's going on in this graph?" section, by helping student participants to notice and wonder about data visualizations.

4 METHOD

4.1 Participant recruitment and interview procedures

In December 2019, our research team coordinated with the Community Editor at the NYT to interview several people involved with content moderation to learn about the steps that they commonly take before, during, and after an online discussion at the NYT website. While the Community Team is responsible for all discussions at the NYT website, many of the regular blogs and learning initiatives take additional community-specific moderation steps.

The Community Editor forwarded information about our study to their team and seven people volunteered to participate. Specifically, the participants include five members of the Community Team and the lead facilitators for the NYT Learning Network "What's going on in this graph?" weekly section. All of the participants have extensive experience as moderators at the NYT. Participants on the Community Team described themselves as lifelong journalists—e.g., being former web producers, reporters, and editors—with a minimum of 6 years and maximum of 12 years moderating discussion at the NYT website. The facilitators at the Learning Network are also former classroom teachers.

Nearly all of the interviews were conducted by video call (one by telephone) during non-working times and each interview lasted 60-90 minutes. All participants provided IRB approved consent to participate in the study. The interview protocol involved a semi-structured set of questions intended to learn about the daily routines involved with moderating user-generated content (from their perspective). The interview questions focused on steps that moderators take when (1) preparing to moderate an article, (2) evaluating the relevance of comments to an article, (3) evaluating the information presented in a comment, (4) looking for comments to elevate in the discussion, and (5) interacting with participants in the discussion. There were no preconditions for the interview

topics; however, we were not permitted to review the official moderation guidelines as these are business sensitive (which is common for commercial platforms [34]).

4.2 Analytic memo-writing analysis procedures

The research team conducted multiple rounds of analytic memo writing [6] based on their compiled notes from each interview. Analytic memo-writing is generally used as a technique in qualitative research, which assists researchers in the process of compiling raw data into abstract themes. In this study, the focus for the analytic memo-writing was to code for the activities related to moderating user-generated content. Four of the paper authors participated in multiple rounds of analytic memo-writing and discussion to identify major and minor themes. This analysis procedure resulted in 21 tasks involved with preparing to moderate (5.1), elevating comments and critique (5.2), rejecting comments (5.3), fact-checking claims (5.4), and synthesizing discussion (5.5).

We chose to present the findings largely in the words of the people who we interviewed. Statements from the interviews are reported with the following unique identifiers: the Community Team are assigned IDs P1-P5 and the Learning Network facilitators are assigned IDs P6 & P7. Each section includes a summary of the tasks associated with each moderation routine.

5 FINDINGS

Our analysis highlights key activities involved with moderating an online discussion that can keep pace with the news cycle. The activities illustrate a range of simple content organization, such as reviewing reader-flagged comments, as well as more complex content organization, such as selecting *NYT Picks* (section 5.2.1) and producing a *Reader Roundup* (section 5.5.1).

5.1 Preparing to moderate

Moderators on the Community Team take shifts managing the first ~24 hours of online discussion about articles, which means that a different member of the team might open, maintain, and close discussion about an article. Moderators prepare to open an article for discussion by reading it to get a sense of what might be on-and-off topic, to “trouble shoot what commenters may be saying about an article” [P5]. However, the time available to prep is limited, “especially during busy days, I’ll probably just read like the top three paragraphs of a story because that’s where all the information is” [P2]. To keep up, moderators “have to speed read them, you know to just get the gist of the story very quickly” [P5].

5.1.1 Moderators get familiar with existing comments and commenters. There are often comments as soon as the discussion opens. As a moderator described, most news stories are open for 24 hours of discussion, “so there’s an urgency to get that point of view [comment] in for sure” [P2]. But, early comments may not be on-topic, “when you open an article for comments and within a minute or two you’ve got 30 of them, you know those people did not read the article” [P5]. The volume of commenting at NYT Blogs, like the *Well Blog*, is a little different:

“We don’t close blog stories [...] so, a lot of people that are just doing Google searches find an article about breast cancer at the Well Blog and they write a comment, but the discussion is like eight years old” [P2].

When moderators shift onto an article that has been open for a while, they read the article, but also spend time learning about the current state of the discussion. The Community Team uses several collaborative technologies to coordinate their efforts around all of the open discussions (e.g., Slack, Google Sheets). “We have spreadsheets to keep track of what’s been open for comment, how long it’s supposed to be open, when it’s supposed to be closed” [P5]. “We share assignments through Slack and Google Docs—so there’s about 10 articles open for comment right now and 10 or

12 in the Opinions—we divvy those up and people get going” [P3]. The Editor and Deputy Editor for the Community Team are responsible for assigning moderators to articles, but the team is in constant contact through Slack, so that they can reorganize their effort as needed.

Moderators also familiarize themselves with any existing comments when shifting onto an open article, “[I will] go through the flags to identify anything horrible, then read the *Reader Picks*, and think about possible comments to highlight as NYT Picks” [P4]. “[Inappropriate content] flags give us a little tap on the shoulder to maybe have another look at a comment, see why six users independently chose to flag it” [P5]. These preparatory steps help moderators to become acclimated to the state of a discussion, before they start moderating new comments in the queue.

Several shared that, as a moderator you begin to recognize regulars in the discussion, “we did have maybe five or six really hardcore conservative people who always had great things to say and who were never up-voted because you know, it’s *The NYT*, so I would always just look for them first” [P2]. As the Community Team manages discussion across many news desks, moderators may gain insights into a regular commenter’s interests, “it would make me chuckle, one of our commenters is very opinionated, almost grating, but to see them post in the [NYT] *Cooking Section*, you see this very different perspective on them as a person” [P4]. Moderators familiarize themselves with existing comments, in part, to see who has already weighed in on the article.

“I look forward to reading comments from specific readers. I have these images of full people based on what they write, a lot of them use their full names and locations, a lot of them don’t, but I feel like they are whole in a way because of their writing” [P4].

5.1.2 Moderators anticipate possible discussion threads. Multiple moderators will manage a high volume discussion: “there could be three or four moderators working on one article, if it was big enough” [P5]. The volume of comments during a shift is hard to predict. Reflecting on one weekend in late January, “we knew the Grammy’s were coming, but we certainly didn’t know that Kobe Bryant was going to die in a helicopter crash, and we had a little bit of advance warning on the Bolton story and that they wanted comments on it” [P5]. Moderators also use social media to monitor breaking news about a story, “if I’m on Twitter then I know that *this* is happening” [P2]. Comments in the discussion can also trigger updates to an article, “like the Kobe Bryant story, people [in the online discussion] said we glossed over his #MeToo moments” [P5].

To prepare for the weekly “What’s going on in this graph?” activity at the Learning Network, moderators will read the news article(s) used in the activity as well as any discussion comments related to an article:

“I read almost all of those [comments]. Really. Because I want to know what people were thinking about and sometimes they add information that the *Times* hadn’t exactly said, so I’ll go check out those sources [hyperlinks] [...] I want to be prepared for somebody saying unexpected things” [P7].

As moderators prepare to open and maintain the conversation about an article, they are trying to anticipate how each discussion might unfold during their shift by reading the article, reflecting on it’s current state, and familiarizing themselves with comments already in the thread, but also by listening for updates to the story.

5.1.3 Summary. Table 1 presents a review of tasks moderators routinely perform when preparing for a discussion. All of the moderation tasks are manual, rather than automated. The Community Team uses collaborative systems to coordinate their efforts, such as Slack, but after distributing the daily assignments, each moderator prepares for their assigned discussion on their own. Moderators read each news article to get a sense of what might be on-and-off topic in the discussion and follow social media to monitor for breaking news.

Task	Description
<i>Applied to all online discussion at NYTimes.com</i>	
1. Review the article topics	Get a sense of what might be on-and-off topic
2. Review existing comments	Get a sense of the discussion so far, by reviewing <i>Reader Picks</i> and comments flagged as <i>inappropriate content</i>
3. Look for frequent users	Pay attention to the commenters who have and who have not yet commented on an article
4. Watch for breaking news	Listen for updates to the story, so as to maintain a sense of what is on-and-off topic during the news cycle
<i>Specific to Learning Network</i>	
5. Collect relevant references	Review past discussions about similar topics to generate a list of related sources to prompt discussion

Table 1. Summary of the tasks involved with preparing to moderate an article discussion. The Community Team and Learning Network are presented separately to highlight how moderators use existing discussion about an article to prepare for a “What’s going on in this graph?” discussion.

5.2 Elevating comments and critique

Moderators watch for comments that add *substance* to the discussion. “Substantial comments” may include opinion, but also clearly describe the reasoning associated with the opinion, whether that reasoning is based in evidence, personal experiences, or beliefs. Like a good discussion prompt [60, 65], substantial comments also nudge the conversation about a topic forward.

“If you feel like borders should be open and everyone should be allowed to come in the country, then just saying that is okay, [but] if you justify why you feel that way through personal experience or a policy or expanding on the viewpoint a little, that is what makes it a substantial comment [that is] more likely to be a springboard for discussion” [P2].

Moderators may not agree with the views expressed by each comment, “even if it’s something that you completely find offensive and disagreeable, if they are putting it in a way that meets our guidelines and they are saying why they feel the way they feel, then we have to go with it” [P1]. Comments that add substance might also be elevated by a moderator—promoted in the discussion or forwarded within the news outlet, to reporters or editors, for instance.

5.2.1 Moderators select representative comments to reflect a range of discussion topics. There are two ways that comments are elevated in the discussion: *Reader Picks* and *NYT Picks*. *Reader Picks* are recommended by readers, while *NYT Picks* are selected by the Community Team. “*Reader Picks* are automatic, so we try to recommend *NYT Picks* that are not already *Reader Picks*” [P4]. *NYT Picks* are a way to showcase the expertise in the commenting audience, “the comments that really add something into the story that’s not necessarily in the article, so could be like the cultural context [or] industry experience, a lot of our readers are experts in their own fields” [P2].

“Sometimes it has to do with an interesting voice, to give you an example: When there were wildfires in California that just kept going, [...] we started running articles about how people were preparing to evacuate [...] we got a comment from someone who was a Director of a Historical Society in Australia [and] archivists have their own procedures about how to do this [i.e., preparing to evacuate]” [P3].

Moderators try to curate the *NYT Picks*, so that the selected comments reflect a range of topics in the discussion, “if you have an article that has 500 comments, we try to raise 15, 20, 25 comments

[as NYT Picks] that represent the range of the conversation, you know, there's only so many people who want to read 500 comments" [P3]. Representing a range of discussion in the NYT Picks does not mean giving equal attention to each view, "if there are 8 million yae and 4 nay, I might not have 4 and 4, the point is looking for people with new views of the argument" [P4]. The Community Team puts a lot of care into selecting NYT Picks, which is why they are often added to the discussion in bulk to present multiple perspectives.

"It looks bad if you put up just one point of view, sitting there solo, while searching for others. Especially in inflammatory pieces. So best practice is to identify multiple points-of-view and add them at the same time to avoid looking like one point-of-view is weighted over another" [P4].

While Reader Picks signal the popular comments, NYT Picks are a collection of substantial comments that reflect a range of topics related to an article.

5.2.2 Moderators watch for comments that critique the reporting. Commenting about an article is also a way for NYT readers to communicate with the newspaper. "In many respects, we are the front line for just about any kind of critique, problem, issue having to do with the paper as a whole. We sort of are the venting place, in many ways" [P1]. As they are looking for comments to elevate in the discussion, moderators also look for comments to promote to reporters, editors, and other teams within the NYT organization.

Some comments offer feedback that could improve an article. For example, "there has been a big push toward more visual aids in our stories, so that's why they [news desks] ask us to sort of keep tabs on what people are saying about data visualizations, and if there are mistakes [...] when there's something skewed, you know, we hear about it first and can pass along the message" [P1]. When a lot of comments raise the same critique, moderators raise these concerns with editors, "sometimes with a story, the people at the [NYT] Reader Center² will do a quick interview with the editor: *Hey you know dozens of commenters complained, can you give us a little bit of insight as to why you chose this headline, so that we can tell our readers?*" [P1]. This type of inter-organizational coordination is one of the ways that online discussion moderation at a news website differs substantially from community moderation elsewhere.

While commenters' often want to communicate directly with reporters [20, 42, 78], it may not be feasible for reporters. Moderators help to filter through the comments, "we try to encourage writers and columnists to respond to really good questions" [P4], "we really do encourage the reporters to answer questions [...] because that's just another way for readers to feel seen and heard" [P1]. However, journalists are still figuring out how to interact with commenters in a productive way, so "our [Community] Team has not always been seen as something that needs to be involved with the journalism, and if journalists are going to interact with the material [comments] it's from their initiative" [P3]. As raised in related work, directly corresponding with commenters can expose a reporters personal opinions, which could lead to accusations of bias about the reporter as well as the news outlet [20, 41].

5.2.3 Summary. Table 2 presents a summary of the moderation tasks involved with elevating content. While all of the tasks are performed manually, each include some routine procedures, such as selecting *NYT Picks* by searching for the most substantial comments from each of the major topics in a discussion. Each of the tasks are also performed with low transparency, but can yield highly visible content (e.g., a response for a journalist). Moderators coordinate with each other via Slack before elevating some content. Moderators encourage journalists to participate by identifying good questions based on comments in the discussion.

²NYT Reader Center: <https://www.nytc.com/press/introducing-the-reader-center/>

Task	Description
<i>Applied to all online discussion at NYTimes.com</i>	
6. Select NYT Picks	Look for multiple points-of-view, highlighting comments that are substantial, but not necessarily the most popular
7. Watch for critique	Pay attention to comments that raise questions about article content (e.g., visual aids, graphs, data analysis)
8. Look for good questions	Pay attention to comments that raise questions that might warrant a response from a reporter or editor

Table 2. Summary of the tasks involved with elevating comments in an article discussion. The moderation tasks in this section are primarily relevant to the Community Team, as they apply generally across NYTimes.com online discussions.

5.3 Rejecting comments

Identifying substantial comments to elevate, either in the discussion or to people within the news outlet (e.g., reporters, editors), is challenging because moderators need to sift through many comments that may not meet the standards for commenting (see Section 3 for information about the NYT guidelines). In this section, we review some of the nuance that makes rejecting comments a tricky task for humans and computers alike.

5.3.1 Recognizing attacks that pass through automated moderation. Since 2017, the Community Team has partnered with Google on an open-source system, called “Moderator”, that applies machine learning to automatically classify the *toxicity* of comments, so as to assist moderators in their process of filtering through comments [25]. “The new moderation tool with the AI component, that is by far the most important development in terms of what we can do because we can just handle so much more volume” [P3].

All of the moderators we spoke with through the interviews felt strongly that the system has improved their performance and reach, as the technology has enabled the Community Team to open online commenting on many more articles. However, moderators occasionally need feedback about why the system has flagged a comment (i.e., possible false positives), “so, sometimes *Moderator* will do us a favor and read someone’s comment at 98%, which would mean highly bad, very dangerous, maybe it had you know a cuss word, it could be as simple as that, but if it didn’t, you’re sort of there drumming your fingers and saying: Well?” [P5].

While moderators feel that the Moderator system does a good job identifying toxic content, some trolling behaviors get through (i.e., false negatives). “We get a lot of people who, I mean it’s childish but the software doesn’t pick up on this, but they’ll change one word and then repost” [P2] or “very often multiple trolls will try to publish the same comment, you know five times, thinking that we reject one, but then the four others [trolls] live on to do their trolling work” [P5]. Moderators can only spend a few minutes addressing these behaviors, because there are so many other comments to moderate, “we don’t have a lot of time to go after this so I’d say the average time we spend is probably about four to five minutes [...] it’s detective work” [P5].

Moderators also rely on commenters to flag comments that had been approved, but may actually violate the guidelines. “Last election, a lot of people were complaining about paid commenters and especially on Russia related stories and [...] we were getting a lot of flags saying like this person is paid” [P2]. However, some commenters will also use this function to attack specific viewpoints in the discussion, “[s]ome of the flagging is just advocacy [...] and any comment that mentions a lower poll result for [a specific candidate], they’ll flag that” [P5].

Existing CSCW research has called attention to the coordinated nature of disinformation operations against online discussion; however, actual evidence of coordination in online discussion can be hard to find [79].

5.3.2 Giving some leeway to let the fight play out. It is hard to moderate comments that fall somewhere in between what is and is not acceptable.

“It’s easy to tell what’s on topic and not abusive on the one hand, and we know what’s really great and substantive on the other, but sometimes there’s a *sticky middle*, like something that’s somewhat troublesome [in a comment], but it’s really true: Is it too inflammatory to post?” [P3].

Every moderator we interviewed talked about sharing these “sticky middle” comments in Slack and asking the team, “what do you think about this one?” [P3]. By remaining in constant daily communication, the Community Team is able to adjust with the volume of discussion and regularly re-calibrate about what is and is not publishable.

Other strategies include reading the articles to recognize when profanity may be topically relevant, “for example, when Trump used the term *s*hole countries* [...] we don’t allow profanity, but this was in our own reporting” [P1]. Personal attacks are always rejected, “you know, a lot of people talk about the Senate majority leader as *Moscow Mitch*, but we issue *name calling*” [P2]; however, there is a sticky middle, “it’s okay for someone to say that Trump is a liar if they have something substantive to say, but if that’s all they’re saying, then it’s just abusive” [P3]. A little bit of substance may be the difference between a personal attack and a substantive comment, “sometimes it’s just a word or a sentence that’s unpublishable” [P4].

Moderators are also sensitive to the fact that some topics can elicit strong feelings from commenters. “[C]alling someone a shthead doesn’t get on, but there’s nuance [...] with every questionable comment we ask ourselves, *does this [comment] move the conversation forward?*” [P4]. As another moderator described, it’s recognizing when the divisiveness helps the community to discuss the topic:

“The [rules of] civility will be a little bit different depending on the topic, so if there’s a good fight going on, you don’t want to totally get in the way of it; you just want it to proceed the way it would in the boxing ring, according to rules. You don’t necessarily want the fight to stop, because that’s part of the spirit of having a community” [P5].

5.3.3 Moderators regularly re-calibrate as a team. In order to maintain a sense of what is and is not acceptable, the Community Team “every month re-calibrates as a group,” to discuss the sticky middle, “but moderators are in constant daily communication *and re-calibration* in two Slack channels, one for opinion and another for news, which everyone on our moderation team monitors” [P4]. When a moderator has a question about whether a comment meets the NYT expectations for online discussion, they will post it to these channels for other moderators to respond, “if we need a *ruling*, we reach out to our community manager or assistant editor, who joins the discussion, and if warranted [they] will send an email to the entire team so all moderators are alerted” [P4]. This constant open channel of communication enables the moderation team to remain consistent in their moderation.

Factors related to the discussion context, such as the author and news section (e.g., Cooking Blog, Opinions, Metro), also play into what is and is not publishable.

“[T]here’s a little bit more leeway with blogs because it’s a little bit more informal [...] the author is talking specifically to his/her audience, so if I’m doing a byline story about Trump’s economic plan, I’m going to probably moderate fairly close to the vest,

not going to approve too much where they're speaking directly to the author, [but] with blogs you kind of court that, because the authors want that feedback" [P1].

Moderators rely on each other, but also consider language in the reporting as well as the topic and setting as they look for comments that advance the conversation.

There are many reasons why a comment may be rejected, but it is not often feasible for moderators to offer this feedback to commenters. "In the [NYT] Cooking Section the volume is much smaller, so we respond to every comment that was rejected and tell them why [but] it's from the *Cooking Desk* not me as a moderator, [otherwise] I would have a lot of penpals" [P4]. Moderators at the Learning Network wish that they had the capacity to explain each rejection, "it would be such a teachable opportunity," without an explanation students may wonder, "*what did I do wrong*, and we don't really have any way of communicating with them about that" [P6].

5.3.4 Summary. Table 3 presents a summary of the moderation tasks involved with rejecting user-generated content. The Google *Moderator* system has enabled the Community Team to manage a high volume of comments, but has also changed their workflow from adjudicating each comment in series to an investigative workflow where moderators reject comments above a toxicity threshold, but also search through the system's predictions for potential false positives (e.g., topically relevant profanity) and false negatives (e.g., trolling behaviors). This investigative workflow reflects a semi-automated, but primarily manual moderation process. Moderators also coordinate with each other about these decisions in real-time. However, some blog discussions at the NYT handle rejections in different ways, for example, all comments rejected from the Cooking Section receive a response from the *Cooking Desk*.

Task	Description
<i>Applied to all online discussion at NYTimes.com</i>	
9. Reject "toxic" comments	Review comments by their predicted toxicity level
10. Look for false positives	Review comments with high predicted toxicity, but no clear cause for rejection, e.g., topically relevant profanity
11. Look for false negatives	Review comments with low predicted toxicity, but may include trolling behaviors, e.g., coordinated attacks
12. Re-calibrate with the team	Deliberate with others on the moderation team about "sticky middle" cases, so as to maintain consistency
<i>Specific to some learning initiatives and blogs (e.g., Cooking Section)</i>	
13. Write rejection message	Send a response so that users know why their comment has been rejected

Table 3. Summary of the tasks involved with rejecting comments in an article discussion. The moderation tasks in this section are primarily relevant to the Community Team with the exception of task #13, that is only applicable in some online blogs at the NYTimes.com website.

5.4 Fact-checking claims

Recently moderators have devoted more attention to fact-checking comments, because they are concerned about the risk of amplifying misinformation:

"I hope this comment is going to be okay, that it doesn't attract, you know, 50 other commenters saying that's not true, any idiot would know that [...] we're in an even harder place if we've made that comment a NYT pick" [P5].

The Community Team has taken steps in the moderation process to mitigate this risk, but also relies on NYT readers to recognize when misinformation gets through, “a lot of times our commenters will give us a heads up: *Hey NYT, Joe Blow put in this comment with this link to something that has been completely debunked, here is the real study*—that happens quite often” [P1]. In this way, moderators and the audience collaborate to fact-check the discussion.

While comments on the main site regularly include links to external information, moderators at the Learning Network “What’s going on in this graph?” encourage students to focus their comments on what they notice and wonder about data visualizations presented in the article that is under discussion. By focusing students on evidence in an article that has already been fact-checked through the NYT publishing process, the moderation team actively works to facilitate informed online discussion about data journalism.

5.4.1 The hyperlinks in a comment can be a first clue for moderators. Moderators follow the news, but are not subject-matter experts in every topic, so they rely on their own judgment and search tools to identify misinformation, “we study *Snopes.com* and other debunking sites [to check] whether it’s real or fake” [P5], “[but] we’re not scientists, we’re journalists [...] sometimes we have to Google [search]” [P1]. When possible, the Community Team will assign moderators to stories that match their interests or experience with a news desk. The team will also reach out to reporters, “sometimes we send science questions to [reporter at the science desk]” [P4].

Moderators pay attention to the websites referenced by a comment to spot possible misinformation. Before a comment can be published at *NYTimes.com*, moderators have to verify each hyperlink in the comment to see that, “it works, it’s not offensive, or [misdirection] pointing to a site to buy shoes” [P4]. Moderators watch for comments that link to known conspiracy sites, “a site like a [conspiracy website] can be a very potent first clue for us,” [P5]. While moderators are asked to verify each hyperlink before publishing a comment, this is a source of anxiety. “If it’s a questionable length then I won’t even click on it because I’m afraid of getting you know, spam or some sort of like weird virus” [P2] and “we’re certainly under a lot of pressure from the tech people to be very careful with these links” [P5].

The Community Team tries to keep track of sites to avoid, “with something like climate change, to name a specific topic, we do want to be able to identify those websites that don’t contain any factual information, sites that are just junk science” [P1]. A bright-line indicator of whether content is factual (or not) is if it has been reported by the NYT, “if it’s a lie and our paper has reported on the lie previously, we would not highlight it [as an NYT Pick]” [P4]. When a story is still evolving, it can be challenging for a moderator to catch up to the community, “you don’t want your commenters to know more about the story than a moderator does” [P5].

Moderators do not have time to fact-check most claims, instead they rely on the community for this support, “I’ve gotten baseball stats wrong and someone will say no actually it was this person who won the Triple Crown in 1985 [...] we have a lot of dedicated commenters” [P1].

5.4.2 Communities can adopt evidence-based facilitation practices. Moderators at the Learning Network section “What’s going on in this graph?” facilitate a weekly discussion-based statistical literacy activity for primary and secondary school students. Each week a different NYT article is selected and revised to conceal a “surprise” observation in a graph that students are meant to identify through an online discussion [36].

Moderators help the students into the data by asking them “*what do you notice?*” and “*what do you wonder?*” about the visualization(s) in an article. Moderators describe these prompts as allowing a low-floor, but high-ceiling for responses, “so we invite comments that might be like, *oh I noticed there are dots on the graph*, just really basic [...] then a high ceiling that could have really rigorous interesting sophisticated things to say” [P7].

After sharing what they notice and wonder, moderators typically respond, “what do you see that makes you say that?” or “what else can you find in the data?” [P6], “my objective when I reply is to take the student to the next level [in their learning]” [P7]. In a post about immigration and crime (e.g., burglary, murder), graphs in the article depicted negative relationships with most variables except for immigration and murder, “some kids would really hone in on the murder statistics and ignore everything else [...] concluding that illegal immigrants are causing crime,” while the comments engage with the graph, moderators would respond by encouraging the student to look at the rest of the data, “*oh that’s interesting, what conclusions do you draw from the other graphs?* to kind of expand their view” [P6].

Over time the moderators have observed some students adopt the facilitation techniques in their own responses, “they’re sort of using that [language] to comment on other people’s posts, which is great especially for those questions that can be kind of contentious, like immigration, I noticed several kids comment: *well, what do you see in the graph that makes you say that?*” [P6]. Moderators at the Learning Network promote statistical literacy by prompting students with questions in a way that encourages them into a discussion and to use evidence presented in an article to explain their thoughts about the topic.

5.4.3 Summary. Table 4 presents a summary of the moderation tasks involved with fact-checking. As they prepare for each discussion, moderators develop a general understanding about many topics, which they rely on to spot misinformation. Moderators also rely heavily on their team and community to flag misinformation *ex post*. While moderators on the Community Team refer to themselves as subject-matter generalists, facilitators at the Learning Network have an expertise in classroom teaching and statistics. Learning Network facilitators participate directly to help the students engage with the data-literacy concepts related to each discussion article.

Task	Description
<i>Applied to all online discussion at NYTimes.com</i>	
14. Survey fact-checking sites	Anticipate misinformation by paying attention to fact-checking sites and prior NYT reporting
15. Review every link	Watch for references to known misinformation websites
16. Watch for reader warnings	Pay attention to reader-replies and “Hey Moderator” messages about possible misinformation in the discussion
<i>Specific to Learning Network</i>	
17. Redirect to the article	Promote discussion about facts presented in the article, rather than opinion and misinformation
18. Respond consistently	Promote specific social norms around community fact-checking, by responding with a consistent social pattern

Table 4. Summary of tasks involved with fact-checking. The tasks are presented separately to reflect observed differences in moderation practices among the teams.

5.5 Synthesizing discussion

A dominant undertone throughout our interviews is that moderators on the Community Team are not just rejecting content. Moderators emphasized that they are reporters and the online community forum is their “beat” (or assigned territory) to follow, so that the newspaper does not miss any important topics that emerge from the online discussion. Grimmelmann [40] describe journalists

as occasionally visiting a community for a story, but *community reporting* is a somewhat new role for discussion moderators.

Not all community discussions should be reported. For example, facilitators at the Learning Network are sensitive to the fact that their audience is largely composed of students and teachers. In this context, the *publicness* [71] of online discussion at the news website can actually deter participation. While some student contributions are highlighted at the end of a week long “What’s going on in this graph?” discussion, facilitators encourage the students to use pseudonyms in the discussion to protect their identity.

5.5.1 Reporting about the online community. It is reasonably common for an editor to ask the Community Team to write a roundup of the reader comments for popular stories, “[...] the editor asked us to do a Roundup to you know, kind of summarize what readers were talking about online, so we just started, there’s probably like 10,000 comments that had been published” [P2]. Moderators research and publish Reader Roundups under their own NYT by-line, “if you look up my by-line, you’ll see a whole bunch of pieces that give that roundup of reader comments,” the objective is, “elevating those comments, sort of get the reader voices out there,” [P3].

A moderator may only have a few hours to synthesize comments about an article into a Reader Roundup about the discussion. Reporting about reader comments on an article is very carefully vetted, “I mean journalists in general have this issue but, for us it was always a sticking point, it wouldn’t get published if it wasn’t well-balanced” [P2]. The following are some of the steps moderators take to synthesize the multiple perspectives into a Reader Roundup:

“First we look at the *Reader Picks* that have been up-voted the most [...] so we look at like maybe the top 10 most recommended ones [...] then we take a deeper dive into the actual comments and kind of try to pick out some really well-written points of view that aren’t so popular, [...] outliers that deserve to be shown. [Keyword] search for terms is a huge thing when trying to find those diamonds in the rough” [P2].

While much of a moderators work is behind the scenes, removing and elevating comments in the discussion, their central role also embeds them within the community, as reporters.

5.5.2 Tending to each community at the news website. In addition to Reader Roundups, the Community Team reports on some social activity among commenters, “there was an article that I wrote about two commenters who met in the Metropolitan Diary Section, so they met online, but then actually met up in real life in Paris” [P2]. The Community Team has also published articles about prolific commenters:

“Two [commenters] have past away, and the NYT ran obituaries for them. [Commenter] was a conservative poster, the nemesis of everybody, but there was such an outpouring of support from the community. People still make comments about missing them, like [Commenter] would have loved this” [P4].

Moderators emphasized that some community discussion should not be reported. Privacy has been an important discussion among moderators at the Learning Network, “you know commenting in this really public way is scary for a lot of teachers, it’s scary for a lot of parents, and kids too” [P6]. One concern is that, “student comments may follow them in the future,” another concern is that there may be classroom dynamics at play, “they’re not just on their computers at home, they’re in a class full of other kids and there’s a lot of peer pressure” [P6].

On some sensitive topics, moderators will ask students to post anonymously, “just maybe put her middle initial or name themselves anonymous, and that might not fly on the regular NYT, [but] we’re a little bit more lenient, because we’re trying to help kids into the conversation” [P6]. While being recognized in the discussion is valuable for many of the most active commenters at the

NYT website [56], the default-settings for *publicness* [71] in the discussion system can also deter community participation.

5.5.3 Summary. Table 5 presents a summary of the moderation tasks involved with synthesizing user-generated content. The moderation tasks in this section involve complex organizations of user-generated content. Each of the tasks are performed manually, though moderators use keyword search and filtering systems to look for comments to include in a Reader Roundup. While the process of writing a Reader Roundup and interviewing community members tends to be low in transparency they yield highly visible products, including stand-alone articles. It is important to note that the NYTimes.com website hosts online discussion at many blogs and learning initiatives in addition to news articles, which may have different expectations for privacy.

Task	Description
<i>Applied to all online discussion at NYTimes.com</i>	
19. Write a <i>Reader Roundup</i>	Summarize the discussion to present multiple points-of-view about a topic, not just the popular perspectives
20. Talk with the community	Make contact with regular commenters to learn about their personal stories related to the community
21. Know when not to report	Pay attention to the privacy norms of each community

Table 5. Summary of the tasks involved with synthesizing comments in an article discussion. Tasks are presented primarily for the Community Team, as they are generally applicable to online discussions across the NYTimes.com website.

6 DISCUSSION

Our analysis revealed five routine activities that moderators perform in order to keep pace with the news cycle: preparing to moderate, elevating comments, rejecting comments, fact-checking claims, and synthesizing discussion. The findings illustrate the multiple roles that moderators play, both to foster online discussion and to support subsequent reporting. First, moderators carefully monitor the online discussion for low-quality content, such as comments that are off-topic or toxic, but also pay attention to emerging topics in the discussion, to stay ahead of the news cycle. Second, while monitoring the online discussion, moderators also facilitate an audience review of the published news (a “secondary gatekeeping”) by routing critiques about the reporting to the right people within the news outlet. Third, moderators come to recognize regular contributors as part of a community, and moderators communicate their observations and insights through community reporting.

6.1 Moderators regulate problematic content and investigate content worth elevating

Our findings capture how the news cycle surrounding a story plays into moderating discussion at a news website. As the news cycle does not stop when an article is published, moderators continuously update their understanding about what is on-topic and factually accurate. These practices demonstrate how moderation at a news website is more than content regulation, it also involves investigating content in the discussion. Moderators will search for valuable perspectives and then fact-check the claims made by a comment before elevating it in the discussion or within the NYT organization—e.g., to get a journalist’s response. Moderators also study the multiple perspectives that emerge around an article in order to synthesize key points into a Reader Roundup.

Our analysis found that most of the investigative work associated with elevating high quality content is performed by hand.

Moderators need automated and semi-automated support to investigate discussion content. There have been considerable advancements in systems used to regulate problematic content (e.g., *Automoderator* [45], *Crosmod* [16], *Digital Juries* [26], Google's *Moderator AI* [25]). Future research should advance systems that assist moderators in surfacing high quality comments (e.g., *CommentIQ* [64]) and understanding the multiple perspectives in a discussion (e.g., *Arkose* [62]). Aitamurto [1] proposed that journalists might leverage crowdsourcing systems, such as those used to coordinate creative writing tasks among large distributed groups of people [27], to “[...] tap into the collective intelligence of the crowds and channel that to their articles” [1, pg. 189].

However, many news outlets, and online communities that facilitate discussion around news, may not have the resources to facilitate the type of content investigation practiced at the NYT. There could be benefits to moderation teams across multiple news organizations collaborating around content investigation, particularly for breaking news [4, 43, 54]. As the news cycle does not stop, well-resourced news outlets could help to bootstrap content investigation at low-resourced news outlets by automatically sharing and updating insights about perspective diversity, popular discussion topics, and examples of misinformation as they emerge in discussion.

6.2 Moderators facilitate a secondary gatekeeping of news articles

News outlets carefully scrutinize published content through a process called gatekeeping (see section 2.4). Our analysis found that during the first 24 hours of online discussion about an article, moderators will watch for comments that provide useful feedback about the reporting. After identifying this feedback, moderators route it through various channels within the news organization, so that the feedback can be evaluated and responded to properly. This process of managing feedback from an online discussion about published articles reflects a “secondary gatekeeping” process, where the news outlet has an opportunity to engage with questions, critique, and leads through updates to the reporting. As an example, moderators shared that the graphics and data teams like to know if and when people struggle to make sense of data visualizations in the article. Moderators also watch for opportunities to involve reporters by responding to questions generated by the online community.

The gatekeeping roles played by moderators imply a somewhat unique position within the NYT organization. At other commercial platforms, moderators often remain hidden [13] and have minimal (if any) impact on their organization [34, 68]. NYT moderators are not recognizable “leaders” like volunteer moderators in many long-standing online communities [22, 58, 86], but they do facilitate communication that connects stakeholders, such as reporters, to conversations at the news website. This secondary gatekeeping can even help reporters to spot errors, clarify points, and respond to breaking news. In this way, moderators at a news outlet may have a dual responsibility for content in the discussion as well as the inter-organizational coordination that happens in response to critiques, questions, and leads surfaced by the discussion.

At many news outlets, there is a clear divide between the newsroom and online discussion about articles [20, 41]. New technologies could help to bridge this gap. For example, with the exception of some sub-communities, like NYT Cooking, moderators do not have the capacity to respond to users whose comments were rejected. Not communicating with people about their rejected comments could be a missed opportunity. Learning more about the people who post unacceptable content could help moderation teams to develop strategies that address these comments *en masse* and that help the newsroom to reach these readers through reporting. Paying attention to the rejection pile may also help newsrooms recognize concerning, yet newsworthy trends. In these ways, technologies that enable moderators to coordinate responses and generate insights from

rejections could add real value to a news outlet. Moderation systems that help news outlets to realize value from a secondary gatekeeping of articles could transform the roles played by moderators, from “custodians” of user-generated content [34, 68] to community reporters.

6.3 Moderators foster online community by reporting about the community

Our analysis shows that moderators, just like the original reporter, operate like journalists, only here they are *reporting from the community beat*. Like other forms of journalism, authoring a Reader Roundup involves overcoming several challenges, such as source anonymity, article balance, and veracity. Moderators have to navigate journalistic standards when synthesizing comments into a Reader Roundup, while also thinking about how the writing might add to subsequent discussions about a newsworthy topic. These activities and considerations related to community reporting are all part of the routine work that NYT moderators perform on a daily basis to advance their mission of promoting civility in public discourse [61].

Like a reporter, moderators also pay attention to the people who frequent the news website discussions, even recognizing familiar profiles by their written voice and where they tend to comment (e.g., Opinion pieces, NYT Cooking). Personal stories that emerge, as moderators get to know people in the community, may turn into published articles (e.g., *A Tribute to a Prolific Times Commenter* [84], *Meet The New York Times’s Super-Commenters* [56], *Widows on the Pain of Dining Alone* [32]). Many people post anonymously in online discussion forums [67], but these and other anecdotes about how online discussion can add value to a person’s life may help to promote a general sense of community at a news website.

Our findings offer examples of how moderators promote civility at the NYT website, but discussion about an NYT article happens on many other platforms. Future research should investigate how the insights generated by moderators at a news website might be *shared* and *updated* through subsequent conversation throughout a discussion ecosystem. Building on existing ideas about cross-community moderation [16], when articles published by the NYT are shared—e.g., through social media, online forums, and group messaging—future systems might enable the recipients to gain access to a synthesis of the insights generated during the first 24 hours of discussion at the publisher’s website. Similar to how *Crossmod* [16] attempts to bootstrap rejection decisions for low resourced online communities, news outlets could help to seed informed discussion about articles they publish by offering examples of popular topics, valued perspectives, and misinformation that emerged during earlier discussions. The process of coordinating insights generated through moderation across multiple communities might also provide news outlets with a means to monitor how discussions develop beyond the publisher’s website.

6.4 Limitations

This paper extends knowledge about online content moderation by highlighting challenges associated with moderating online discussion during a fast-paced news cycle. We draw on the Grimmelmann [40] definitions for online community moderation to discuss limitations associated with our study. First, Grimmelmann [40] characterizes four primary members of an online community, including owners of the infrastructure, moderators of the community, and authors and readers of the content; this paper focuses exclusively on the routines of moderators. Additionally, our study protocol asked the moderator participants to retrospectively reflect on their routines, rather than observe them in action (as in a Cognitive Walkthrough [77]). Second, Grimmelmann [40] proposes that moderation routines can be implemented in a variety of ways (e.g., automatically versus manually). Due to issues of access and time we were not able to study the Google Moderator AI [25], interview members of the online community who regularly flag inappropriate content, or speak with reporters who have joined in the online discussion.

Finally, our study is limited to the views of seven moderators from one news outlet. A broader investigation could analyze data that draws from multiple moderation teams and makes comparisons across multiple platforms. Our findings might serve as a design probe to learn about how other moderation teams have introduced similar and different practices. A comparative study could also help to create a taxonomy of moderation support systems used by news outlets. While the NYT platform is custom built, other news websites use commercial and open-source discussion platforms (e.g., *Coral*, *Discourse*) and third-party comment hosting services (e.g., *Disqus*). The NYT is a unique case in many respects, but the findings offer inspiration and points of comparison for other news outlets and online communities that host discussion about the news.

7 CONCLUSION

In this paper, we explore discussion moderation practices at the New York Times website. Our interview participants include five members of the NYT Community Team as well as lead facilitators of the “What’s going on in this graph?” section at the NYT Learning Network. The interview findings present illustrative examples of how moderators manage online discussion about published content at the news website. Our analysis identified a series of moderation routines to limit low-quality content and promote high-quality content in online discussions. Due to the fast-paced nature of the news cycle, moderators are constantly updating their own mental model of topics and facts that are relevant and accurate. This investigative work is primarily performed manually, which suggests ample opportunities to enhance the capabilities of moderators with some level of automated or semi-automated support.

Moderators also play a “secondary gatekeeping” role within the community by facilitating opportunities for NYT stakeholders, like reporters, to engage with questions, critiques, and leads surfaced by the discussion. The central role that moderators play in this inter-organizational coordination at the NYT is considerably different from the situation of moderators at other commercial and community platforms. While it may not be feasible for other news outlets and online communities to adopt the NYT practices, we discuss a cross-community model of moderation, where insights generated during the moderation process are *shared* along with an article and then *updated* through continued discussion of the article.

ACKNOWLEDGMENTS

We thank all of the Community Moderators at the New York Times for sharing their time and thoughts with us. This research would not have been possible without support from Marcia Loughran, Community Editor at the New York Times, and Natalie Proulx, Senior Staff Editor at The Learning Network. The paper improved tremendously through the review cycle and is much stronger thanks to the thoughtful contributions of expert reviewers. This research was funded by the National Science Foundation (grant #2009003). Additionally, we want to thank Eric P. S. Baumer, Molly Q Feldman, and Stephen MacNeil for sage advice on the research.

REFERENCES

- [1] Tanja Aitamurto. 2016. Crowdsourcing in Open Journalism: Benefits, challenges, and value creation. *The Routledge Companion to Digital Journalism Studies* (11 2016), 185–193. <https://doi.org/10.4324/9781315713793-19>
- [2] Kimberley Allison and Kay Bussey. 2020. Communal Quirks and Circlejerks: A Taxonomy of Processes Contributing to Insularity in Online Communities. In *International AAAI Conference on Web and Social Media*, Vol. 14. 12–23.
- [3] Ashley A. Anderson, Dominique Brossard, Dietram A. Scheufele, Michael A. Xenos, and Peter Ladwig. 2014. The “Nasty Effect”: Online Incivility and Risk Perceptions of Emerging Technologies. *Journal of Computer-Mediated Communication* 19, 3 (4 2014), 373–387. <https://doi.org/10.1111/jcc4.12009>
- [4] Cynthia Andrews, Elodie Fichet, Yuwei Ding, Emma S. Spiro, and Kate Starbird. 2016. Keeping up with the tweetdashians: The impact of ‘official’ accounts on online rumoring. In *Proceedings of the ACM Conference on Computer*

- Supported Cooperative Work*, CSCW, Vol. 27. Association for Computing Machinery, New York, NY, USA, 452–465. <https://doi.org/10.1145/2818048.2819986>
- [5] Michael Barthel. 2019. *State of the News Media: Trends and Facts on Newspapers*. Technical Report. Pew Research Center. <https://www.journalism.org/fact-sheet/newspapers/>
 - [6] Melanie Birks, Ysanne Chapman, and Karen Francis. 2008. Memoing in qualitative research. *Journal of Research in Nursing* 13, 1 (1 2008), 68–75. <https://doi.org/10.1177/1744987107081254>
 - [7] Robin Blom, Serena Carpenter, Brian J. Bowe, and Ryan Lange. 2014. Frequent Contributors Within U.S. Newspaper Comment Forums. *American Behavioral Scientist* 58, 10 (9 2014), 1314–1328. <https://doi.org/10.1177/0002764214527094>
 - [8] David Bogen. 1983. The Origins of Freedom of Speech and Press. *Maryland Law Review* 42 (1983).
 - [9] John Bowers and Jonathan Zittrain. 2020. Answering Impossible Questions: Content Governance in an Age of Disinformation. *Harvard Kennedy School Misinformation Review* (1 2020). <https://doi.org/10.37016/mr-2020-005>
 - [10] Arthur S. Brisbane. 2010. Readers With Plenty to Say. *The New York Times* (12 2010).
 - [11] Byron Calame. 2006. Voices of Readers Online, Windows on the Newsroom. *The New York Times* (12 2006).
 - [12] Lily Canter. 2013. The Misconception of Online Commenting Threads. *Journalism Practice* 7, 5 (10 2013), 604–619. <https://doi.org/10.1080/17512786.2012.740172>
 - [13] Elinor Carmi. 2019. The Hidden Listeners: Regulating the Line from Telephone Operators to Content Moderators. *International Journal of Communication* 13, 1 (1 2019), 440–458.
 - [14] Alissa Centivany and Bobby Glushko. 2016. “Popcorn Tastes Good”: Participatory Policymaking and Reddit’s “AM-Ageddon”. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI16)*, San Jose, CA.
 - [15] Stevie Chancellor, Jessica Pater, Trustin Clear, Eric Gilbert, and Munmun De Choudhury. 2016. Thyghgapp: Instagram content moderation and lexical variation in Pro-Eating disorder communities. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, Vol. 27. Association for Computing Machinery, New York, NY, USA, 1201–1213. <https://doi.org/10.1145/2818048.2819963>
 - [16] Eshwar Chandrasekharan, Chaitrali Gandhi, Matthew Wortley Mustelier, and Eric Gilbert. 2019. CrossMod: A cross-community learning-based system to assist reddit moderators. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (11 2019), 1–30. <https://doi.org/10.1145/3359276>
 - [17] Kevin Coe, Kate Kenski, and Stephen A. Rains. 2014. Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments. *Journal of Communication* 64, 4 (8 2014), 658–679. <https://doi.org/10.1111/jcom.12104>
 - [18] Stephen Coleman and John Gøtze. 2001. *Bowling together: Online public engagement in policy deliberation*. Hansard Society London.
 - [19] Mark Deuze, Axel Bruns, and Christoph Neuberger. 2007. Preparing for an age of participatory news. *Journalism Practice* 1, 3 (2007), 322–338. <https://doi.org/10.1080/17512780701504864>
 - [20] Nicholas Diakopoulos and Mor Naaman. 2011. Topicality, time, and sentiment in online news comments. In *CHI’11 Extended Abstracts on Human Factors in Computing Systems*. 1405–1410.
 - [21] Jill P Dimond, Michaelanne Dye, Daphne LaRose, and Amy S Bruckman. 2013. Hollaback!: the role of storytelling online in a social movement organization. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 477–490.
 - [22] Bryan Dosono and Bryan Semaan. 2019. Moderation practices as emotional labor in sustaining online communities: The case of AAPI identity work on reddit. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300372>
 - [23] Justin Ellis. 2015. What happened after 7 news sites got rid of reader comments. *Nieman Journalism Lab* (9 2015).
 - [24] Bassey Etim. 2014. A Comment’s Path to Publication. *The New York Times* (4 2014).
 - [25] Bassey Etim. 2017. The Times Sharply Increases Articles Open for Comments, Using Google’s Technology. *The New York Times* (6 2017).
 - [26] Jenny Fan and Amy X. Zhang. 2020. Digital Juries: A Civics-Oriented Approach to Platform Governance. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376293>
 - [27] Molly Q Feldman and Brian James McInnis. 2021. How We Write with Crowds. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (1 2021), 1–31. <https://doi.org/10.1145/3432928>
 - [28] Casey Fiesler, Joshua McCann, Kyle Frye, and Jed R Brubaker. 2018. Reddit Rules! Characterizing an Ecosystem of Governance. In *Twelfth International AAAI Conference on Web and Social Media*. <https://aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17898>
 - [29] Casey Fiesler, Shannon Morrison, R. Benjamin Shapiro, and Amy S. Bruckman. 2017. Growing Their Own. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW ’17*. ACM Press, New York, New York, USA, 1375–1386. <https://doi.org/10.1145/2998181.2998210>
 - [30] Rolf Fredheim, Alfred Moore, and John Naughton. 2015. Anonymity and Online Commenting: The Broken Windows Effect and the End of Drive-by Commenting. In *Proceedings of the ACM Web Science Conference*. ACM, Oxford, United

- Kingdom, 11. <https://doi.org/10.1145/2786451.2786459>
- [31] Lena Frischlich, Svenja Boberg, and Thorsten Quandt. 2019. Comment Sections as Targets of Dark Participation? Journalists' Evaluation and Moderation of Deviant User Comments. *Journalism Studies* 20, 14 (10 2019), 2014–2033. <https://doi.org/10.1080/1461670X.2018.1556320>
- [32] Aidan Gardiner. 2019. 'After His Death, I Didn't Cook Anymore': Widows on the Pain of Dining Alone. *The New York Times* (11 2019).
- [33] R. Stuart Geiger and David Ribes. 2010. The work of sustaining order in wikipedia: the banning of a vandal. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. ACM, 117–126. <https://doi.org/10.1145/1718918.1718941>
- [34] Tarleton Gillespie. 2018. *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press. 1–288 pages.
- [35] Tarleton Gillespie. 2018. Platforms are not intermediaries. *Georgetown Law Technology Review* 2 (2018), 198–216.
- [36] Michael Gonchar and Katherine Schulten. 2017. Announcing a New Monthly Feature: What's Going On in This Graph? *The New York Times* (9 2017).
- [37] Robert Gorwa. 2019. What is platform governance? *Information Communication and Society* 22, 6 (5 2019), 854–871. <https://doi.org/10.1080/1369118X.2019.1573914>
- [38] Robert Gorwa, Reuben Binns, and Christian Katzenbach. 2020. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society* 7, 1 (1 2020), 205395171989794. <https://doi.org/10.1177/2053951719897945>
- [39] Elizabeth Grieco. 2020. *U.S. newspapers have shed half of their newsroom employees since 2008*. Technical Report. 1 pages. <https://www.pewresearch.org/fact-tank/2020/04/20/u-s-newsroom-employment-has-dropped-by-a-quarter-since-2008/>
- [40] James Grimmelmann. 2015. The Virtues of Moderation. *Yale Journal of Law and Technology* 17, 1 (2015), 42–109. <https://heinonline.org/HOL/P?h=hein.journals/yjolt17&i=42>
- [41] Alfred Hermida and Neil Thurman. 2008. A Clash of Cultures. *Journalism Practice* 2, 3 (10 2008), 343–356. <https://doi.org/10.1080/17512780802054538>
- [42] Sanne Hille and Piet Bakker. 2014. Engaging the Social News User. *Journalism Practice* 8, 5 (9 2014), 563–572. <https://doi.org/10.1080/17512786.2014.899758>
- [43] Mengdie Hu, Shixia Liu, Furu Wei, Yingcai Wu, John Stasko, and Kwan Liu Ma. 2012. Breaking news on Twitter. In *Conference on Human Factors in Computing Systems - Proceedings*. ACM Press, New York, New York, USA, 2751–2754. <https://doi.org/10.1145/2207676.2208672>
- [44] Elizabeth Jensen. 2016. NPR Website To Get Rid Of Comments. *National Public Radio (NPR)* (8 2016).
- [45] Shagun Jhaver, Iris Birman, Eric Gilbert, and Amy Bruckman. 2019. Human-machine collaboration for content regulation: The case of reddit automoderator. *ACM Transactions on Computer-Human Interaction* 26, 5 (7 2019). <https://doi.org/10.1145/3338243>
- [46] Shagun Jhaver, Amy Bruckman, and Eric Gilbert. 2019. Does transparency in moderation really matter?: User behavior after content removal explanations on reddit. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (11 2019), 1–27. <https://doi.org/10.1145/3359252>
- [47] Inyoung Kang. 2019. Where Does California's Homeless Population Come From? *The New York Times* (11 2019).
- [48] Kate Klonick. 2017. The New Governors: The People, Rules, and Processes Governing Online Speech by Kate Klonick. *Harvard Law Review* 131, 1598 (2017). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2937985
- [49] Robert E. Kraut, Paul Resnick, Sara Kiesler, Moira Burke, Yan Chen, Niki Kittur, Joseph Konstan, Yuqing Ren, and John Riedl. 2012. *Building successful online communities: Evidence-based social design*. MIT Press.
- [50] Suzanne LaBarre. 2013. Why We're Shutting Off Our Comments. *Popular Science* (9 2013).
- [51] Cliff Lampe and Paul Resnick. 2004. Slash (dot) and burn: distributed moderation in a large online conversation space. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 543–550.
- [52] Cliff Lampe, Paul Zube, Jusil Lee, Chul Hyun Park, and Erik Johnston. 2014. Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Government Information Quarterly* 31, 2 (4 2014), 317–326. <https://doi.org/10.1016/j.giq.2013.11.005>
- [53] Alex Leavitt and Joshua Clark. 2014. Upvoting hurricane sandy: Event-based news production processes on a social news site. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 1495–1504. <https://doi.org/10.1145/2556288.2557140>
- [54] Alex Leavitt and John J. Robinson. 2017. The role of information visibility in network gatekeeping: Information aggregation on reddit during crisis events. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*. Association for Computing Machinery, New York, NY, USA, 1246–1261. <https://doi.org/10.1145/2998181.2998299>

- [55] Joseph Lichterman. 2016. 20 years ago today, NYTimes.com debuted “on-line” on the web. *Nieman Journalism Lab* (1 2016).
- [56] Kat Long. 2017. Meet The New York Times’s Super-Commenters. *The New York Times* (11 2017).
- [57] Lena Mamykina, Drashko Nakikj, and Noemie Elhadad. 2015. Collective Sensemaking in Online Health Forums. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. ACM Press, New York, New York, USA, 3217–3226. <https://doi.org/10.1145/2702123.2702566>
- [58] J. Nathan Matias. 2019. The Civic Labor of Volunteer Moderators Online. *Social Media + Society* 5, 2 (4 2019), 205630511983677. <https://doi.org/10.1177/2056305119836778>
- [59] J. Nathan Matias and Merry Mou. 2018. CivilServant: Community-led experiments in platform governance. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2018-April. Association for Computing Machinery, New York, New York, USA, 1–13. <https://doi.org/10.1145/3173574.3173583>
- [60] Alfred Moore. 2012. Following from the front: Theorizing deliberative facilitation. *Critical Policy Studies* 6, 2 (7 2012), 146–162. <https://doi.org/10.1080/19460171.2012.689735>
- [61] Heather Moore. 2008. Talk to the Newsroom: Community Editor. *The New York Times* (8 2008).
- [62] Kevin K. Nam and Mark S. Ackerman. 2007. Arkose: Reusing informal information from online discussions. In *GROUP’07 - Proceedings of the 2007 International ACM Conference on Supporting Group Work*. ACM Press, New York, New York, USA, 137–146. <https://doi.org/10.1145/1316624.1316644>
- [63] NYT Learning Network. 2019. Teaching Resources for Bringing the World Into Your Classroom. *The New York Times* (9 2019).
- [64] Deokgun Park, Simranjit Sachar, Nicholas Diakopoulos, and Niklas Elmquist. 2016. Supporting comment moderators in identifying high quality online news comments. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, 1114–1125. <https://doi.org/10.1145/2858036.2858389>
- [65] Walter C. Parker. 2006. Public Discourses in Schools: Purposes, Problems, Possibilities. *Educational Researcher* 35, 8 (11 2006), 11–18. <https://doi.org/10.3102/0013189X035008011>
- [66] Kate Phillips. 2007. The Caucus Blog: Caucus Chatter. *The New York Times* (6 2007).
- [67] Emma Pierson. 2015. Outnumbered but well-spoken: Female Commenters in the New York times. In *CSCW 2015 - Proceedings of the 2015 ACM International Conference on Computer-Supported Cooperative Work and Social Computing*. Association for Computing Machinery, Inc, New York, NY, USA, 1201–1213. <https://doi.org/10.1145/2675133.2675134>
- [68] Sarah Roberts. 2016. Commercial Content Moderation: Digital Laborers’ Dirty Work. *Media Studies Publications* (1 2016). <https://ir.lib.uwo.ca/commpub/12>
- [69] Joseph Seering. 2020. Reconsidering Self-Moderation. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (10 2020), 1–28. <https://doi.org/10.1145/3415178>
- [70] Joseph Seering, Tony Wang, Jina Yoon, and Geoff Kaufman. 2019. Moderator engagement and community development in the age of algorithms. *New Media & Society* 21, 7 (7 2019), 1417–1443. <https://doi.org/10.1177/1461444818821316>
- [71] Bryan C. Semaan, Heather Faucett, Scott P. Robertson, Misa Maruyama, and Sara Douglas. 2015. Designing Political Deliberation Environments to Support Interactions in the Public Sphere. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI ’15*. ACM Press, New York, New York, USA, 3167–3176. <https://doi.org/10.1145/2702123.2702403>
- [72] Pamela J. Shoemaker, Martin Eichholz, Eunyi Kim, and Brenda Wrigley. 2001. Individual and Routine Forces in Gatekeeping. *Journalism & Mass Communication Quarterly* 78, 2 (6 2001), 233–246. <https://doi.org/10.1177/107769900107800202>
- [73] Pamela J. Shoemaker, James W. Tankard, and Dominic L. Lasorsa. 2003. *How to build social science theories*. Sage Publications.
- [74] Jane B. Singer. 2010. Quality Control: Perceived effects of user-generated content on newsroom norms, values and routines. *Journalism Practice* 4, 2 (4 2010), 127–142. <https://doi.org/10.1080/17512780903391979>
- [75] Jane B. Singer and Ian Ashman. 2009. “Comment Is Free, but Facts Are Sacred”: User-generated Content and Ethical Constructs at the Guardian. *Journal of Mass Media Ethics* 24, 1 (3 2009), 3–21. <https://doi.org/10.1080/08900520802644345>
- [76] Jonathan Smith. 2016. We’re Getting Rid of Comments on VICE.com. *Vice News* (12 2016).
- [77] Rick Spencer. 2000. The streamlined cognitive walkthrough method, working around social constraints encountered in a software development company. In *Conference on Human Factors in Computing Systems - Proceedings*. ACM Press, New York, New York, USA, 353–359. <https://doi.org/10.1145/332040.332456>
- [78] Nina Springer, Ines Engelmann, and Christian Pfaffinger. 2015. User comments: motives and inhibitors to write and read. *Information, Communication & Society* 18, 7 (7 2015), 798–815. <https://doi.org/10.1080/1369118X.2014.997268>
- [79] Kate Starbird, Ahmer Arif, and Tom Wilson. 2019. Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (11 2019). <https://doi.org/10.1145/3359229>

- [80] Natalie Jomini Stroud, Joshua M. Scacco, Ashley Muddiman, and Alexander L. Curry. 2015. Changing Deliberative Norms on News Organizations' Facebook Sites. *Journal of Computer-Mediated Communication* 20, 2 (3 2015), 188–203. <https://doi.org/10.1111/jcc4.12104>
- [81] The New York Times Company. 2012. *The New York Times Company Reports 2012 Third-Quarter Results*. Technical Report. The New York Times Company, New York.
- [82] The New York Times Company. 2021. *The New York Times Company Reports 2020 Fourth-Quarter and Full-Year Results*. Technical Report. The New York Times Company, New York.
- [83] Jeremy W. Peters. 2011. The Times announces digital subscription plan. *The New York Times* (3 2011).
- [84] Nancy Wartik. 2018. A Tribute to a Prolific Times Commenter. *The New York Times* (12 2018).
- [85] David Manning White. 1950. The "Gate Keeper": A Case Study in the Selection of News. *Journalism Quarterly* 27, 4 (9 1950), 383–390. <https://doi.org/10.1177/107769905002700403>
- [86] Donghee Yvette Wohn. 2019. Volunteer moderators in twitch micro communities: How they get involved, the roles they play, and the emotional labor they experience. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300390>
- [87] Bingjie Yu, Joseph Seering, Katta Spiel, and Leon Watts. 2020. "taking care of a fruit tree": Nurturing as a layer of concern in online community moderation. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3334480.3383009>

Received January 2021; revised April 2021; accepted May 2021