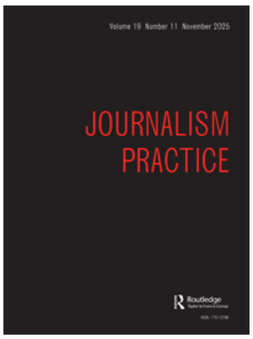

Routledge
Taylor & Francis Group

Journalism Practice

ISSN: 1751-2786 (Print) 1751-2794 (Online) Journal homepage: www.tandfonline.com/journals/rjop20

Journalists "Attagged": The @-tag as a Bonding Tool for "Supercharged Critical Publics"

Marloes Geboers, Tomás Dodds, Mark Boukes & Eirliani Abdul Rahman

To cite this article: Marloes Geboers, Tomás Dodds, Mark Boukes & Eirliani Abdul Rahman (05 Aug 2025): Journalists "Attagged": The @-tag as a Bonding Tool for "Supercharged Critical Publics", Journalism Practice, DOI: 10.1080/17512786.2025.2539338

To link to this article: https://doi.org/10.1080/17512786.2025.2539338

Published online: 05 Aug 2025.

JOURNALISM PRACTICE
https://doi.org/10.1080/17512786.2025.2539338

Journalists "Attagged": The @-tag as a Bonding Tool for "Supercharged Critical Publics"

Marloes Geboers[a], Tomás Dodds[b], Mark Boukes[c] and Eirliani Abdul Rahman[d]

[a]Media Studies, University of Amsterdam, Amsterdam, the Netherlands; [b]Journalism and Mass Communication, University of Wisconsin–Madison, Madison, WI, USA; [c]Communication Science, University of Amsterdam, Amsterdam, the Netherlands; [d]Collaborative on Global Children's Issues, Georgetown University, Washington, DC, USA

ABSTRACT

Social media's attention-based economy and subsequent design, allegedly, spur mistrust and attacks against media workers. X (formerly "Twitter") facilitates a discursive climate that is divisive and polarized. Within this context, we present research on platform-afforded tagging practices as infrastructures that enact "subtle" underminings of journalistic authority through piggybacking on benign expressions of press-critique. X does not allow effective "untagging" of oneself from other people's tweets. We showcase how this creates a material addressability, amplifying harassment. The case zooms in on the repetitive questioning of whether or not BBC journalist Laura Kuenssberg was "at the party", alluding to Partygate, a scandal about meetings held at 10 Downing Street during Covid-19 lockdowns. Repurposing linkages between hashtagged and @-tagged tweets allowed us to map a network of critique that assembles a community of so-called "supercharged critical thinkers". While the "repetitive drum of suspicion" might seem benign, the tagged tweets, asking whether Kuenssberg was at the party, are embedded within a networked ecology of vitriol and misogyny through hashtags that connect creators to a wider network of distrust. While one can exert some control over tagging by segmenting one's audience, such options are too rudimentary for journalists who have to fulfill a public role.

ARTICLE HISTORY
Received 2 December 2024
Accepted 21 July 2025

KEYWORDS
Media-critique; networked harassment; platform affordances; digital methods; context collapse; platform design

Introduction

Academic research into online violence and harassment of journalists often focuses on severe cases of hate speech, such as those that emerge in cases of so-called "dogpiling", where a coordinated attack materializes in "a barrage of threats, slurs, insults, and other abusive tactics".[1] A report by UNESCO (Posetti et al. 2022) outlines the so-called chilling effects of social media harassment on journalists' work. News workers share cases of them withdrawing from reporting on sensitive issues, or even leaving the profession

CONTACT Marloes Geboers m.a.geboers@uva.nl

altogether. Harassment turned out to be gendered with, across various studies worldwide, 73% of women reporting having experienced online harassment (Cinnamon 2024; Posetti et al. 2022).[2] Solutions center on tools for journalists' safety, which can, for example, safeguard a journalist's personal details. However, once a dogpile begins, controlling the damage becomes nearly impossible (Cinnamon 2024). The focus of many tools is on developing mitigation or damage control. However, such tools result from blind spots for the role that the design of the platform and its subsequent tagging architecture plays in the sustenance of violent behavior. The design of communicative features of the platform, such as the options a user has for effectively untagging themselves at scale, sustains networked entanglements between hate and violence on the one hand and more benign expressions of legitimate critique on the other hand. Our case study will show the intricate connections between critique and violence, as well as the role of the communicative tagging architecture of X in maintaining such connections. By conducting a network analysis between replies (@ tagged tweets) and hashtags, as well as a temporal analysis of persistent and recurring accounts engaging in harassment, we render legible how harassment unfolds both as an extension of wider collective distrust and as a product of the design of social media platforms such as X. Visualizing the materialities of a journalist's digital environment, lays bare the imperceptible routes of connection that platforms afford and that entangle suspicion with hateful comments in platform-specific ways (Hendrickx and Opgenhaffen 2024). We aim to enhance our understanding of digital environmentality as well as platform-specific logics that deeply influence how journalists engage with news, news commentary, and audiences on social media.

Cases of dogpiling and their "chilling effects" no doubt require urgent scholarly attention; however, the wider networked environments that host severe attacks require our attention as well. What is not sufficiently understood is the role of social media's use design in channeling harassment. Arguably, the user affordances of many social platforms, such as hashtags and @-tags, create a material addressability of journalists that is difficult to escape. Once someone is dogpiled on X (formerly known as Twitter), this person can, in principle, untag themselves from these posts; however, one would need to do that post per post. Another "option" is to allow tagging only to accounts that the journalist follows. However, in light of the public function of journalists, as well as the professional dependencies that many journalists have on platforms, this is merely a "faux choice". We stress that the current premise underlying safety measures for journalists is insufficient by default as the communicative design of platforms prioritizes engagement over user empowerment and well-being (Christin et al. 2024).

Given the difficulties of untagging from social media posts, it is also important to explore how seemingly benign or even positive and constructive forms of critique directed at journalists can quickly get connected to potential dogpilers and to more vile kinds of content. As soon as someone is addressed through the @-tag with a critical remark, this remark can be mimicked, recontextualized, and connected to much more toxic content. Adding hashtags such as #scummedia renders a question directed at a journalist, such as a sincere "Were you at the party?", less innocent. In this study, we investigate how legitimate critique on platforms, such as X, serves as an easy vector for harassment, collapsing legitimate critique with violent behavior. We investigate a case study that displays particular dynamics of press critique that are exemplary of social

media collective effervescence channeled by a critical question directed at BBC journalist Laura Kuenssberg: "Were you at the party?" referencing so-called Partygate. The latter was a scandal in the United Kingdom about (festive) government staff gatherings during the COVID-19 pandemic when health restrictions were in place.[3]

We gathered @-tagged replies to the prominent UK journalists between November 25 and December 26, 2021, when the scandal was brought to light by the Daily Mirror. The temporalities of hashtags were mapped to analyze the overall discourse and the pertinence of hashtags that connected tweets to allegations of bias and partiality. Most notable were hashtags that emerged from the replies to the then political editor of the British Broadcasting Corporation (BBC), Laura Kuenssberg. From a close reading of more than 118,727 tweets, we found an interesting pattern wherein users (N = 1,510 posts) would ask Laura Kuenssberg the question "Were you at the party?" or some close variation of this question. We then charted how this seemingly benign question is connected or networked to downright toxic or violent messages through situating the questions in their wider hashtagged environment using the method of network analysis.

Literature Review

In the literature, the role of journalists as gatekeepers of news both online and offline (Bro and Wallberg 2014; Janowitz 1975; Shoemaker 1991; Shoemaker, Vos, and Reese 2008) is well researched. Less clear is their role in spotlighting specific content and authors of posts (Zhou et al. 2023) and what happens in the amplification process. Journalists are influenced by the commercial logic of social media and by influential actors and collectivities, who intentionally or unintentionally destabilize their position (Nieborg and Poell 2018; Van Dijck, Poell, and Waal 2018). It is, unfortunately, reasonable to expect that high-profile journalists' very public role would invite harassment and attacks of the sort expected of public personalities. In the attention economy, journalists may be pushed to maximize engagement, which may expose them to harassment (Wischerath et al. 2022). Online violence against journalists has also increased (Barton and Storm 2014; Miller 2021; Waisbord 2020a). Although digital violence or defamation of journalists does not equate to media criticism, it occurs beyond the end of the spectrum of critique.

The decline of public confidence in news media (Hanitzsch, Van Dalen, and Steindl 2018) has multiple societal causes, not all of them related to mediatization. However, it is safe to say that social media not only made journalists more visible but also their errors and dependencies more noticeable (Davies 2009; Marwick 2021). Following Phillips and Milner (2021), we consider social platforms as part of an ecology in which journalistic outlets are forced to compete with influencers and campaigners while being subject to algorithmic amplification or subluxation. Journalists must navigate this online ecology while, to a degree, being undermined in various ways and by different actors (Ong and Tapsell 2022), who, as our study shows, are not always intentionally contributing to wider harassment but are nonetheless paving the way for harassment through merely outing critique. Influencers or campaigners explicitly challenge legacy media by destabilizing their role in shaping social discourse, questioning production processes, or even initiating hate campaigns against journalists (Cheruiyot 2022; Pintak, Bowe, and Albright 2021). Some of these campaigns push narratives of perceived personal fallacies against media workers or accusations based on normative critiques using traditional journalistic concepts, such as objectivity and

neutrality (Miller 2021; Waisbord 2020a). Within this influence ecology, journalists' legitimacy is continuously contested, even, or especially so, when they try to lean into social media logic through elements like clickbait (Petre 2021).

The infrastructure of press criticism coming from news audiences was revolutionized by the affordances of the Web 2.0, which enlarged the realms of critique from "letters to the editor" to an interactive space hosting a much more comprehensive range of discourses and voices (Cheruiyot 2022). Praised for its democratizing potential, the web brought hopes of a participatory culture in journalism, positioning citizens as the "fifth estate" (Hayes 2008). All of these tools and practices are continuously being modified or newly created, making it rather difficult to define what constitutes "more established" cultural techniques or practices.

However, the steep power asymmetries built into contemporary social media platforms have hampered the realization of this democratic ideal (Leavitt 2013). Hopes of citizen-empowered journalism have been dashed as the web scaled up and got monopolized by tech companies that thrive on "sticky news" and upheaval; their business model is built on the exploitation of free affective labor (Berry and Sobieraj 2013; Rathje, Van Bavel, and van der Linden 2021). Like most social media platforms, X's business model is based on an attention economy that places emotion at the forefront of journalistic practices. Berry and Sobieraj (2013) argue for considering the economic underpinnings of what they dub an "outrage industry".

X is market-driven and dependent on the stickiness of content circulating on its platform. In other words, content must first and foremost grab users' eyeballs and prolong the time they spend on the platform. What makes them stick? In the context of problematic and (hyper-)partisan news media, Berger and Milkman's study into viral news content (2012) defines what activates users to share content by assessing the emotive components of more or less shared content. They found that content that evokes high-arousal positive emotions (awe) or high-arousal negative emotions (anger or anxiety) is more engaging. Content that produces low arousal, or deactivating, emotions (e.g., sadness) is less so. Thus, outrage is engaging, which sheds light on how news on X becomes "engaging" when it is: "meant to cause outrage, cue partisan emotions, and get clicks (i.e., make money)". Even though almost all journalistic artifacts hold emotionality to some extent (Wahl-Jorgensen 2019), the affordances of X facilitate a discursive climate that is more extreme, divisive, and polarized (Shepherd et al. 2015) compared to traditional news spaces. This dynamic constitutes a "turbulent space" in Waisbord's (2020b) words, wherein normative and personal attacks have become rife to the point where several news organizations have left the platform altogether (see the announcement from The Guardian in November 2024 to refrain from posting to X). However, due to the specific dependencies on social platforms for distribution and dissemination (i.e., revenue), journalists continue to be present on social media in active ways, using diverging strategies to cope with and respond to criticisms.

The Analytical Value of Centering Platform Affordances

Platform-embedded methods of connection—i.e., the hashtag and the @-tag in particular—and the communicative architecture they construct hold analytical value for the study of networked harassment. Adopting a "digital methods approach" (Rogers 2018), we

repurposed the networked infrastructure that emerges from hashtagging and @-tagging practices, which organize platform users around a shared sentiment or commonality vis-á-vis "the press". Tagging as a founding feature of the participatory web was hailed for its democratizing potential in terms of information organization. Nicola Bozzi (2023, 71) frames tagging—in its broader definition—as "an operational form of identity labeling that contributes to naturalizing historically grounded practices of social classification" and, in so doing, can widen or narrow meaning. Introduced in the mid-2000s by the bookmarking site del.icio.us, tags allowed users to share links and label them individually through the use of textual keywords that made them easily searchable and accessible through the website (Bozzi 2020, 72). Functioning as linguistic markers, hashtags can be understood as metadata or labels signifying the intent with which one uploads content to platforms (Zappavigna 2015). As previously mentioned, X does not allow for feasible ways of "untagging" yourself from other people's tweets, resulting in a material addressability (Honey and Herring 2009) that allows anyone to "link up" to anyone else by mentioning their Twitter handle. This function can facilitate forms of harassment (Phillips and Milner 2018), which tends to exert heavier effects on users whose social identity is defined by intersectional markers, including gender, race, or sexual identity/orientation, but also profession (Chen et al. 2020; Miller and Lewis 2022; Posetti et al. 2022).

We explore the various ways in which tagging features are a practice that is central to the exchange of social value. Aligning with news and sharing or associating yourself with news on social platforms is, first and foremost, normatively valuable as an act of democratic participation and as political self-representation in relation to others. Information transmission and perhaps coercion or persuasion are then secondary. Within this social-ritualistic context, hashtags serve to affiliate with or distance oneself from a certain stance or issue (Zappavigna 2015). The @-tag then establishes a communicative connection in the shape of a "reply or a response to". Following the premise that social exchange value exceeds the value of information transmission (Dean 2021; Lewis, Marwick, and Partin 2021), this sought-after connection is first and foremost a performance of belonging or affiliation.

Ritualistic Environments and the Intertwinement of Critique and Hate Speech

Both hashtagging as well as @-tagging can be understood as participatory practices of identity-formation (Bozzi 2023). Within this social context, the study of hashtagging or "@-tagging" journalists can be understood as a networked practice, which can revolve around critiquing the so-called "mainstream press". We understand this critique as being part of an interaction ritual (Collins 2014) that establishes membership to a community that is, in this case, consisting of what boyd once called "supercharged critical publics" in a public talk on misinformation (2018). Phillips and Milner (2021) link the emergence of such publics to the polluted information flows of networked ecologies of influence that co-opt doubt and suspicion; thereby leveraging uncertainty as part of a business model. In adopting the lens of the interaction "ritual" (Collins 2014) we take on a Durkheimian perspective on online communities, where the exchange of arguments in discussions should not be primarily viewed as the rational exchange of thoughts, but also as content that serves the role of symbolic markers of identity and belonging (Törnberg and Törnberg 2024).

Approaching social media commenting practices and tagging through this lens, contentious engagements, such as press critique, can potentially spill over into harassment as part of a contagious process (Song et al. 2022). In such fundamentally social processes of networked bonding, connections are inherently made on the basis of in- and outgroup communication that needs to draw on vernacular language to exclude the outgroup. This need for "exclusive language" leads to an increased interpretive flexibility of hate speech (Ganesh 2018). Many scholars focus on the ways in which memes, irony, and in-jokes, oftentimes ambiguous, make it difficult for moderators to determine when they cross the line into hate speech. What is oftentimes overlooked is that it is not merely the ambiguous nature of borderline memetic behaviors that poses problems, it is also the ways in which blatant hate "thrives and survives" because of its entangled connections to seemingly "innocent" formations of content.

Another point that Ganesh (2018, 46) makes in his work on the "ungovernability of hate speech" centers the weaponization of the Twitter purge of 2017 by those who were deplatformed. The lack of clarity on the selection criteria for banning extremist accounts by Twitter only

> enraged the swarm and mobilized it to encourage users to move to a more extreme space, Gab.ai. In a similar vein, any attempts at censoring or using legal action against these individuals must contend with their response, which identifies any use of power against them as further evidence of the allegedly repressive liberal establishment.

In social media commentary cultures, somewhat similar dispositions against mainstream media have taken hold, and while healthy critique of the press is an inherent part of a functioning democracy in spaces where doubt and uncertainty are commodified, institutions can easily find themselves painted as "the enemy of the people". This is a particular kind of tension that arises in interactive spaces that was aptly foreseen in the findings of a survey by Downes and McMillan (2000) in the early days of the web. Some respondents speculated about interactivity as a threat (to authorial control), empowerment (of audiences or users), and opposition to media industries (as controllers). One of their respondents: "Because I have a voice now, I think the threat is to existing institutions and the old ways of doing things (164)". Conflating institutions with top-down control of narratives seems to have bolstered this suspicion, which now materializes in a widespread presence of conspiracy theories that more often than not also host distrust against legacy media. Indeed, the societal benefits of some of the speech that can be found online often appear unclear; however, unsavory online language cannot uniformly be deemed as simply "bad" as it can serve positive, even necessary, societal and democratic functions.

Evidence from multi-country surveys has demonstrated that using social media as a main source of news reduces citizens' trust in the news (Kalogeropoulos et al. 2019) and increases general mistrust in the news (Park et al. 2020). There are various explanations for this. For example, many people consume the mainstream news on social media; however, while doing so they often do not recognize the journalistic organization that has produced it. So, even if they found it credible or enjoyable, many will not realize that the news stories they had read originated from a regular news medium (Kalogeropoulos, Fletcher, and Nielsen (2018) 2019). Besides this lack of a positive effect, social media users are likely to be exposed to media bias accusations (Strömbäck and Åkerlund

2025) or claims of "fake news" (Meeks 2020). Reading such social media comments may, therefore, negatively influence the public's trust in the media and public opinion toward media or individual journalists (Dobber and Hameleers 2025). Such effects of social media commentary cultures are often found in survey-based studies or ethnographies probing news audiences, but the erosion of trust in legacy news media is also spurred by social platforms' commercial constructs of commodification.

Social media use design is tailored for the optimization of engagement contingent on affective labor and the already mentioned "outrage industry" (Berry and Sobieraj 2013). In this context, Robertson, Del Rosario, and Van Bavel (2024) found how a small group of commenters creates a dynamic they dub "false polarization" and "pluralistic ignorance". The authors link these phenomena to a number of problems, including intergroup hostility, support for authoritarian regimes, and, by extension, we argue, a supercharged cynical form of anti-press hostility. The distorted representation of norms in the online surroundings of hate and harassment—i.e., social media's use design disproportionately "visibilizes" extreme ideas and expressions—more nuanced or moderate takes tend to risk hostility from more extreme ingroup and outgroup members, especially since such hostility has little cost for the aggressor due to the social distance that the online environment affords (Robertson, Del Rosario, and Van Bavel 2024).

The material features of social media environments clearly shape the conditions under which journalists and their audiences interact; this inspired us to shift focus to the materialities of X and map out how platform-afforded connections undergird harassment that collapses critique and hate speech, and that is dominated by in-group and out-group animosities, demarcating a desire to belong to "supercharged critical publics".

Methods

The project commenced with setting up the collection of tweet metadata, based on a compiled query consisting of the social media handle of BBC's political editor, Laura Kuenssberg (i.e.,@bbclaurak).[4] This effectively assembles a tweet using her handle, with or without added hashtags. For data collection, we used the Twitter Research API v2, which was then widely accessible for researchers. Currently, a similar collection could be conducted through browser-based scraping tools such as the DMI Zeeschuimer Firefox extension tool (Peeters 2021). At that time, toward the end of 2021, Kuenssberg covered the allegations of a government party being held during COVID-19 lockdowns. Her reporting on this scandal attracted not merely attention to Partygate, but also registered in tweets conveying distrust against her professional integrity, questioning her impartiality. We decided to hone in on @-tagged tweets at her address, published between November 25 and December 21, 2021. This time frame spans points in time that were relevant both to the unfolding of the scandal and to the parallel critique (spilling over into harassment) aimed at Kuenssberg personally. This dataset comprised a total of 118,727 tweets tagging her account @bbclaurak. The number of tweets, on average, is more than 4.5 thousand a day, demonstrating the momentum that this case attracted on the platform.

The first analytical step was to chart the hashtags present in the tweets that were tagging Kuenssberg (Figure 2), using the DMI 4CAT capture and analysis tool (Peeters and Hagen 2022) that, among many other things, allows for mapping out hashtags and

their frequencies over time. The time span includes Partygate, explaining many of the most frequent hashtags, which also makes sense as Kuenssberg was among the journalists who were actively reporting about this scandal. These hashtags are accompanied by several anti-press hashtags, some pertaining to Kuenssberg specifically. Inspired by the most frequently used hashtag #downingstreetparty in week 49, boasting the highest frequency number, we then proceeded by querying the dataset using "the party" as a search string, and filtering out said tweets.

This led us to detect the substantial repetition of the question Were you at the party Laura? (n = 1,510). We mapped the question-tweets over time, identifying upticks in presence that coincided with news reports relating to the scandal (Figure 3). Alongside this, we mapped the persistence of active users (Figure 4) who were engaging multiple times in asking the same rhetorical question, exemplifying how repetition is a technique for the prolongation and consolidation of media critique in a dynamic environment. From there, a hashtag-mentions network (Figure 5) was made with the Gephi spatialization tool (Bastian, Heymann, and Jacomy 2009). This renders visible how the tweets asking the "party-question" are connected to more toxic harassment.

Findings

The first reporting of Partygate was on 30 November 2021 by the Daily Mirror. Allegedly, Downing Street 10 staff gatherings had taken place during the 2020 Christmas season. This unleashed wide critique of not merely politicians, but also of those reporting on it. On December 1, an account with a significant following (10.3 K followers) reflected on the news by @-tagging Kuenssberg and highlighted how Kuenssberg refused to answer the question of whether she was at the party or not. As mentioned, healthy critique coming from news audiences is an inherent part of a functioning press in democracies; however, within the volatile and highly emotive spaces of X, such tweets quickly tease out responses that not only amplify critique but also connect suspicion and distrust to sheer vitriol and misogyny.

In the time span of three weeks following December 1, Kuenssberg was asked the question of whether she was at the party 1510 times. These questions were embedded in threads that host various textual and visual messages that frame her not only as (potentially) impartial, but that also leverage more vile messages that are networked through the infrastructure of # and @ linkages. Networked connections between critique, bullying, and misogyny emerge from the indexing capacities of hashtags that are used in tandem with the tagging of Kuenssberg's account. In this way, Kuenssberg's persona (@-tagged) collapses with accusations of the Tory party as being corrupt (#torylies, #torylaura). Such conflations also unfold in multimodal trajectories where images appear in reply threads to tagged tweets. Although these multimodal trajectories were not our main focus of analysis, we include some examples for illustration in Figure 1. A returning example is the highly polished image depicting Laura as "Secretary of State Propaganda". Another is a photograph of Kuenssberg and Boris Johnson sitting on a bench, overlaid with text balloons that state how Kuenssberg would ask him how she can be of help.

When mapping hashtag frequencies from November 25 to December 21 (weeks 47–51) (Figure 2), we see how Kuenssberg gets discursively contextualized with hashtags that are anti-establishment (#toryscum, #torycorruption) as well as anti-press

Figure 1. Assembles some examples of tweets targeting Kuenssberg, purposefully leaving out more severe derogatory posts, such as those with highly sexist takes such as the journalist having a "milk shower".

(#scummedia, #defundthebbc). The underlying data is, as mentioned in the previous section, consisting of tweets and quote tweets that tag Laura Kuenssberg's account (@bbclaurak).

The most frequently used hashtag, in tweets tagging Kuenssberg, is #downingstreet-party in week 49. This hashtag assembles many innocent tweets, as Kuenssberg reports quite actively on Partygate. However, in the wake of her professional presence on X, Kuenssberg also receives critique of impartiality and insincerity (#torylaura, N = 13) that extends to her employer (#defundthebbc, N = 69) or even to "the media in general" (#scummedia, N = 105). These point to tweets that instrumentalize the raised suspicion of Kuenssberg being at the party and of her alleged conservative bias, in order to weaponize their wider anti-media and anti-BBC campaigns. From the hashtags #whoattendedthepary and #whowasattheparty (accumulating 299 tweets), we can derive that her name is discursively connected to the party allegations.

In Figure 3 we see how the repetitive, rhetorical asking of the party-question, is rather contingent on events. Spiking intensity occurs around December 8, coinciding with the publishing of a video of a "mock" press conference where staffers are seen joking about a party. Downing Street Press Secretary Allegra Stratton resigned shortly after the video surfaced. While not mentioning Laura Kuenssberg, let alone insinuating her alleged attendance, the appearance of the video does ignite an intensification of suspicion. Kuenssberg's reporting on it, and the further dissemination of her journalistic item, is enough for "critical" X-users to veer up.

When we visualize the accounts involved in the spike of occurrences of the rhetorical question (Figure 3), we can see in Figure 4 that some accounts display "inauthentic" or at

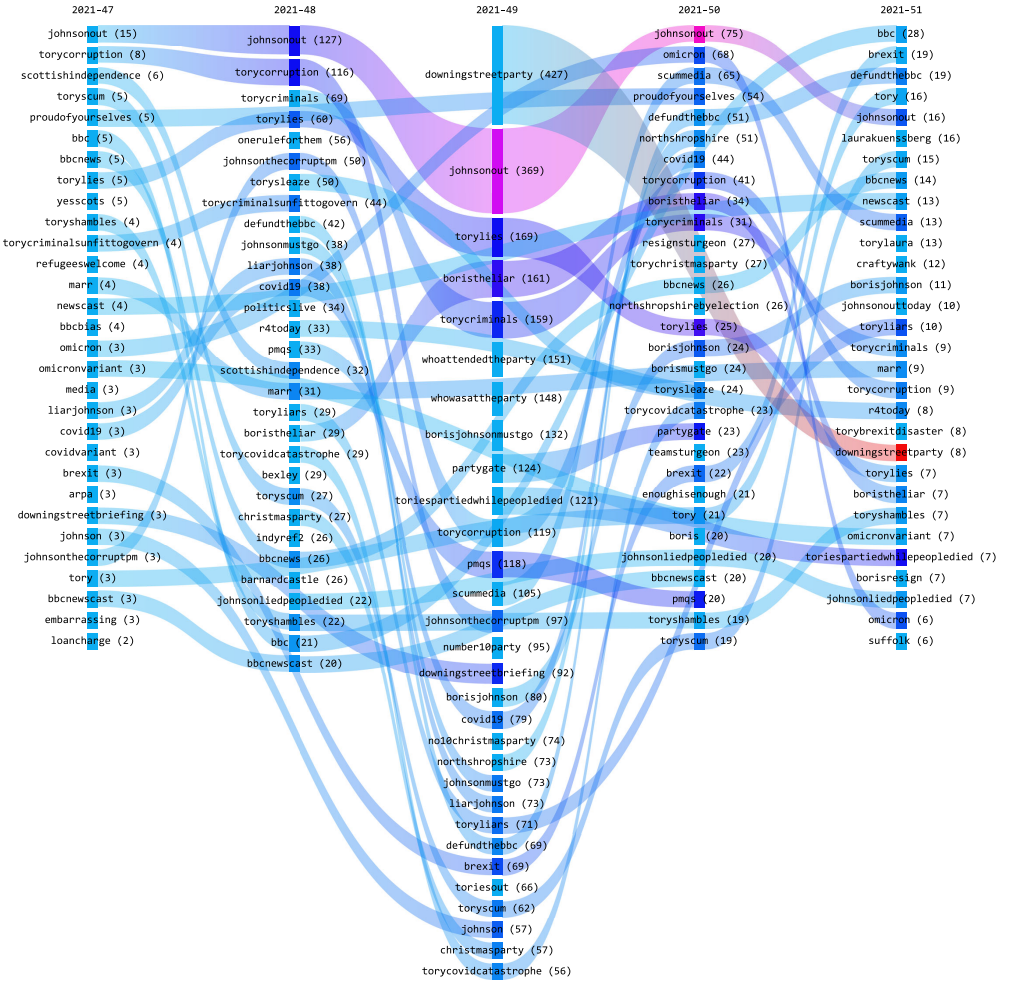

Figure 2. Represents frequent hashtags (with use numbers) over a time span of five weeks, for the 118,727 tweets that @-tag Laura Kuenssberg.

least persistent behavior. The accounts seem not to be automated, but, for example, account 3 (purple in Figure 4) repeated the question 10 times in a timespan of about seven hours. This suggests great arousal, presumably outrage, or perhaps, to a lesser extent, mockery. It can point to a more coordinated organic effort. In this context, it is interesting to note that one of the top three persistent users, as outlined in Figure 4 asks the question seven times and uses the hashtag #whoattendedtheparty in an effort to enhance visibility. At the time of writing, these accounts were still up and active in a seemingly organic manner (not automated).

Guilt by Association

The network analysis in Figure 5 depicts the wider discursive and networked environment wherein the question "Were you at the party Laura?" is embedded. A small sum of these questions intersect with hashtags, such as #torylaura. This shows how the subtle but

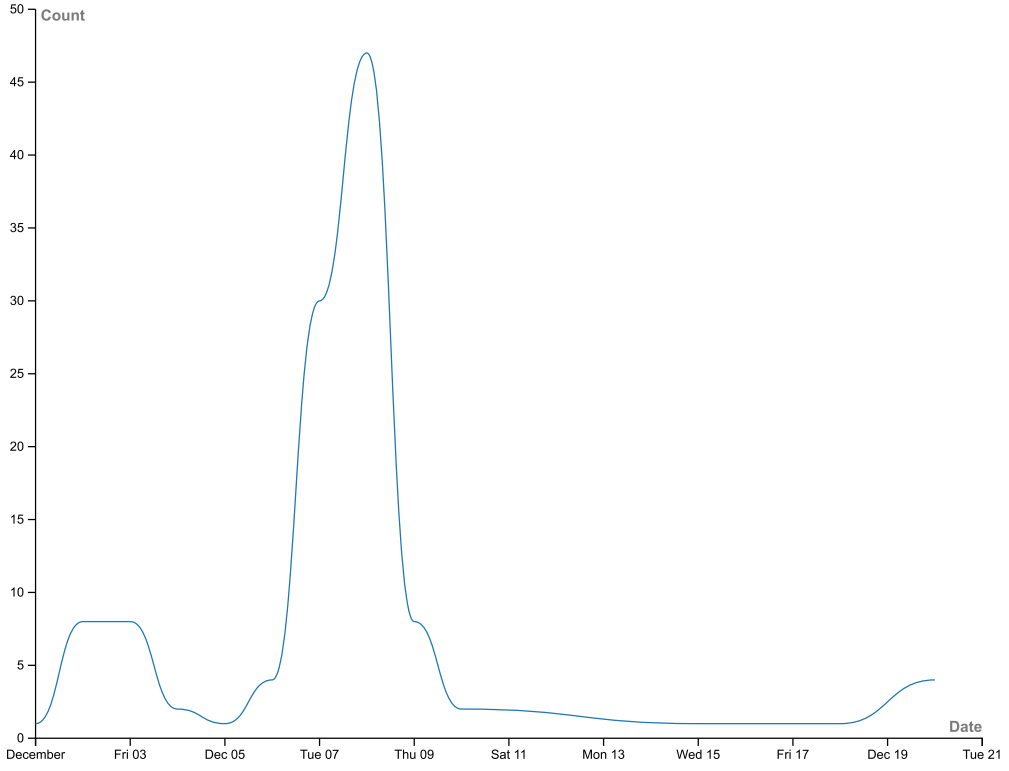

Figure 3. Depicts the temporal distribution of tweets asking the question: Laura, were you at the party?

pervasive repetitions of "critique" in the shape of a seemingly benign question is connected to, and leveraged by wider distrust of legacy media "as mouthpieces of the corrupt establishment". Some examples of these @-replies that collapse the "benign" question with more straightforward harassment, include: "@bbclaurak Laura were you at the Downing Street xmas party on December 18th last year? Just asking for a friend. #torysleaze #oneruleforthem".[5] And, also: "@bbclaurak @PippaCrerar @danbloom1 Shit reporting as per the norm! Your 'source' (and who knows it could have been you at the party) could have leaked this last year but no … as usual silence from the British Bullshit Corporation #BBCBias #ToryScum". And: "@bbclaurak @DoreJayne Hey #ToryLaura were YOU at the party? If you were jilted, even more reason to let us know who was. Obviously #LiarJohnson but which other ministers?" One of the earliest tweets: "@bbclaurak @PippaCrerar @danbloom1 It's not a scoop @bbclaurak if you were at the party! #torylaura, #ToryCriminals, #ToryLies, #ToryCorruption". Note how the hashtags in these tweets collapse the journalist's name and the party where she is supposedly a mouthpiece for.

Reflecting on the entanglements between the question raising doubt about the professional integrity of Kuenssberg on the one hand and its connections to anti-media hashtags partly resonates with what has been conceptualized as context collapse (Marwick and boyd 2011). When journalists get tagged, entanglements with other layers of expression or other digital artifacts such as news items about PartyGate that as standalone artifacts do not evidence nor imply Kuenssberg being involved in the illegal

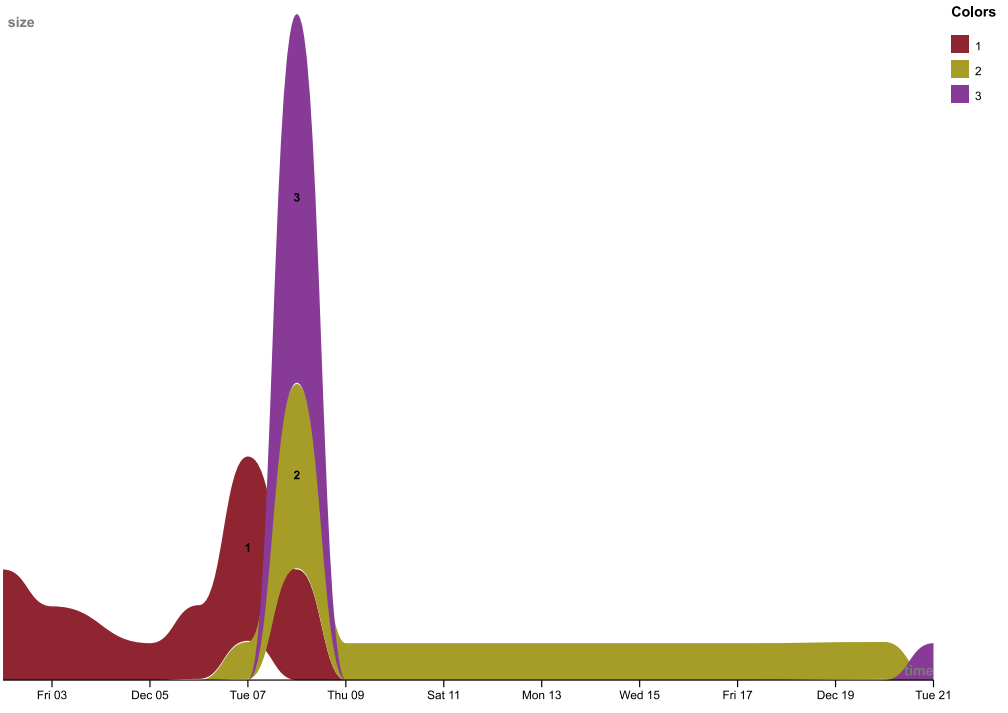

Figure 4. Displays a stream graph where the colors represent the three most persistent users and the extent to which they repeated the question.

parties at Downing Street 10, but that through assembly (Parry 2022) contribute to a ritualistic display of "guilt by association". The rhetorical question and variations thereof recontextualize and weaponize news headlines and items so as to bolster this shared suspicion (Brandtzaeg and Lüders 2018; Phillips and Milner 2021).

Discussion

Our study furthers an understanding of how specific platform affordances, such as (hash)-tagging, facilitate the connection between seemingly benign jokes and healthy public critique on the one hand, and more extreme content on the other hand. From our study, the critical but also largely rhetorical question of "Laura, were you at the party?" found a wider audience through Kuenssberg's handle as @-tag as well as via hashtags. Alongside this networked act of contagious suspicion, it is also a way for individual users to latch onto a perceived crowd that exerts affective allure through its simulation of "being with many". The use of hashtags and the @-tag, thus, opens up "the stage" for anti-press sentiment and enhances wider visibility of public distrust of journalists and the media in general.

In this context, it is important to point out the skewed constituency of active "commenters" on social media. Recent research (Robertson, Del Rosario, and Van Bavel 2024) found that online discussions are dominated by a small, extremely vocal, and non-representative minority. While only 3% of active accounts are toxic, they produce 33% of all content. This minority not only disseminates hateful content, but they also bias meta-perceptions of

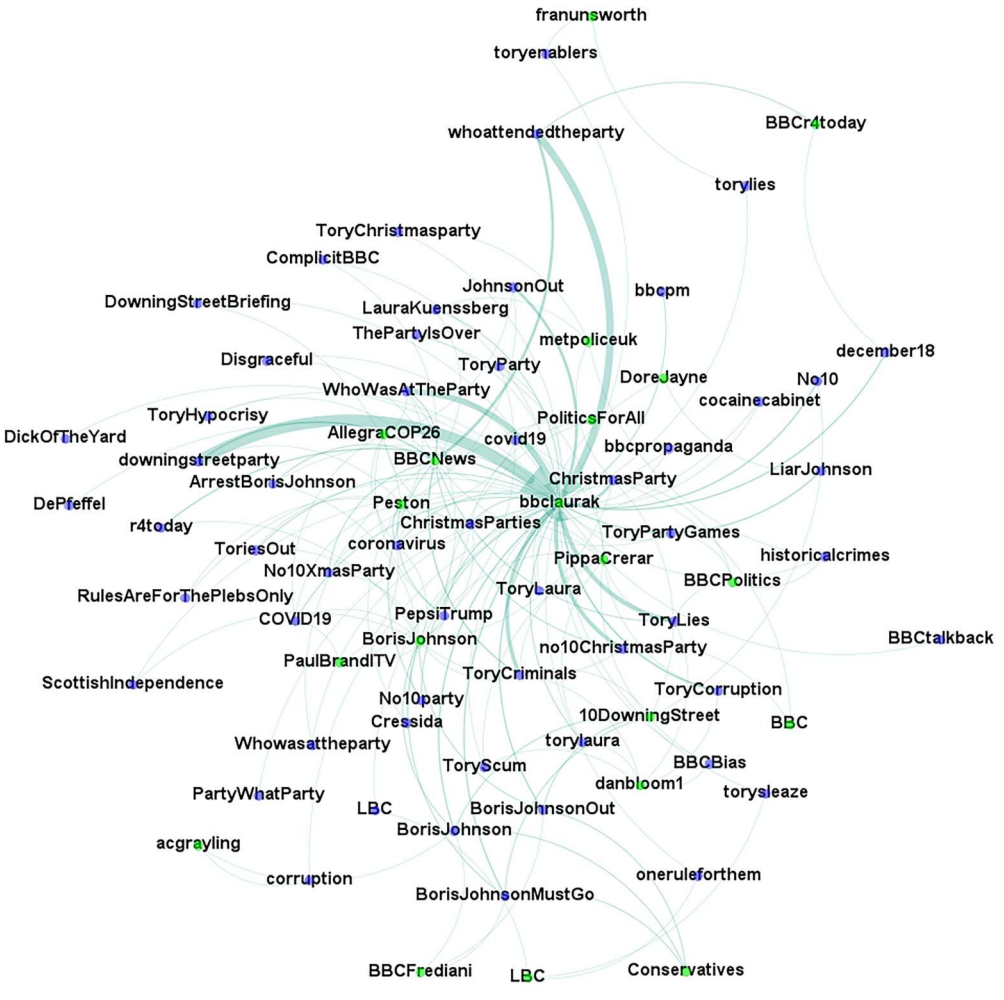

Figure 5. Represents the networked connections (at least two connections) that @-tag the account of Laura Kuenssberg (@bbclaurak at the center of the network). Accounts that are tagging her account are displayed as green nodes and hashtags that were used in these tweets are depicted as blue nodes. The blue nodes clearly include those that emerge from public distrust in news media, such as #toryenablers, #BBCBias, #bbcpropaganda, and #torylaura.

users who passively lurk. These so-called lurkers create distorted representations that disproportionately "visibilize" extreme ideas and expressions. Some of our findings exemplify this; in the time span of seven hours, just one user was responsible for ten party-question tweets.

Tags and their indexing capacities establish an infrastructure for social bonding and boundary work, assembling supercharged critical publics. While seemingly innocent, the party's question-tweets, when they are put on repeat, become pivotal to the entrenchment of particular ideas about a fraudulent press. They are also part of a wider discursive network that connects users to more vile content. Repetitive networked associations between journalists and politicians-under-scrutiny, and between insincere behaviors or characteristics (biased) and particular professions (media), are quintessential

for ideas to solidify. In such a communicative environment, "nuance takes too long", according to Jodi Dean (2021, ix). Ambiguous or nuanced positions struggle to gain attention; anything more intricate than simple rhetorical statements like "Were you at the party Laura, Y/N?" cannot gain traction. Dean (2021, X) summarizes this dynamic as "your emoji defeated my argument". Logical reasoning is, by default, defeated by short messages that travel faster because they conflate and collapse complexities and particularities.

In the context of social media images, Jodi Dean (2021) conceptualized visual content as having transformed from artifacts holding exhibition value to artifacts that are available by way of their commonality. This can be expanded to our case study, wherein a critical question directed at a journalist of the BBC became a rhetorical tool for bonding among people generally cynical about the media. The communicative environment is likewise one of a generalized discursivity that is characterized by multiplicity, repetition, and association; in this case, the association is with a belonging to super critical audiences who "do their own research". In a ritualistic display of belonging, aligning oneself with others who are "also super critical" then becomes a matter of mimicry, where in the case of our study 1,240 distinct accounts engaged in repeating the same question in a time span of three weeks, signaling their belonging to a group who "dares to ask". More or less tactically using the @-tag and several anti-press hashtags, these users establish visibility within a much wider community of supercharged critical thinkers.

Accounting for Ecology

Research into criticism and harassment directed at journalists tends to gloss over the ecological logic of social media, wherein every individual act can culminate in harassment on a wider scale (Phillips and Milner 2021). This phenomenon is contingent on the networking affordances of said platforms and, more specifically, on the design on which these affordances hinge. Within this context, the role of a platform's communicative design and its subsequent consequences for user agency are as yet understudied. Users have options to exert some control over who is tagging you; however, within X, this is merely an option for tagging you in photos. Blocked accounts can still @-tag a journalist, although the journalist themselves cannot see that particular tweet. In short, user options are highly rigid and constrained, leaving journalists with little to no feasible tools to address interactions that spiral into, or connect to, downright toxic attacks. Allowing a user to disable tagging at scale, or preventing blocked accounts from tagging (not merely for photos), would at least dampen the visibility of malign actors and repeated associations that these actors construct between the journalist and certain wider anti-press vernaculars.

Our research exemplifies how hash- and @-tagging not only channels violent language but is also complicit in the perpetuation of "subtle" and ambient modes of undermining journalistic authority (Siapera 2019). It is imperative to understand that even with the most intricate moderation, the networked ecology of platforms (Marwick 2021) makes sure that short routes to toxic content remain intact. The rather innocent @-tagged tweets of our case study exemplify how such tweets are connected to vitriol and misogyny by way of hashtags that connect content and creators to a wider network of public distrust, where the path to conspiracy theories and harassment is short. In other words, people engaging by sharing or participating in bully memes, such as the one in

our case study, intentionally or unintentionally connect themselves and others to misogynist and downright violent content.

In the wider context of mapping and interpreting hate speech, scholarly attention often centers on memes, irony, and in-jokes, as the ambiguous modes of social boundary work online render moderation notoriously difficult (Ganesh 2018). Our approach–centering the networked paths between content formations afforded by various tagging strategies–acknowledges the intertwinement between healthy critique of the press and downright malicious actors whose only aim is to destabilize public trust. Getting a grasp of these platformed dynamics is the first step to more effectively anticipating how particular news items might be weaponized and transformed into tools that generate backlash against journalists. Furthermore, the role of repetition to solidify ideas and the human desire for commonality as a central prerequisite for the dissemination of harassment is pertinent and deserves further scholarly attention.

Notes

1. From the PEN online harassment field manual. Available at: https://onlineharassmentfieldmanual.pen.org/defining-online-harassment-a-glossary-of-terms/.
2. The UNESCO-ICFJ survey (2020) recorded that 73% (n = 456) of women journalists surveyed at the global level who answered a specific question regarding incidence said they had experienced online violence in the course of their work.
3. https://en.wikipedia.org/wiki/Partygate.
4. Initially queried were also the handles (user names) of Owen Jones, Julia Hartley-Brewer, Amelia Gentleman, Paul Brand, Ashley Cowburn, Vicki Young.
5. This tweet was deleted by the time of writing, however, we could still retrieve the caption text as collected metadata through the Twitter V2 API, that was at that time still available.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

References

Barton, A., and H. Storm. 2014. *Violence and Harassment Against Women in the News Media: A Global Picture*. Washington DC: International Women's Media Foundation & International News Safety Institute. https://www.iwmf.org/wp-content/uploads/2018/06/Violence-and-Harassment-against-Women-in-the-News-Media.pdf.

Bastian, M., S. Heymann, and M. Jacomy. 2009. "Gephi: An Open Source Software for Exploring and Manipulating Networks." *Proceedings of the International AAAI Conference on Web and Social Media* 3 (1): 361–362. https://doi.org/10.1609/icwsm.v3i1.13937.

Berger, J., and K. L. Milkman. 2012. "What Makes Online Content Viral?" *Journal of Marketing Research* 49 (2): 192–205. https://doi.org/10.1509/jmr.10.0353.

Berry, J. M., and S. Sobieraj. 2013. *The Outrage Industry: Political Opinion Media and the New Incivility*. Oxford: Oxford University Press.

boyd, d. 2018. "You Think You Want Media Literacy … Do You? Data & Society: Points, Medium." https://medium.com/datasociety-points/you-think-you-want-media-literacy-do-you-7cad6af18ec2.

Bozzi, N. 2020. "Tagging Aesthetics: From Networks to Cultural Avatars." *A Peer-Reviewed Journal About* 9 (1): 70–81. https://doi.org/10.7146/aprja.v9i1.121490.

Bozzi, N. 2023. "Machine Vision and Tagging Aesthetics: Assembling Socio-technical Subjects through New Media Art." *Open Library of Humanities* 9 (2): 1–25. https://doi.org/10.16995/olh.10023.

Brandtzaeg, P. B., and M. Lüders. 2018. "Time Collapse in Social Media: Extending the Context Collapse." *Social Media + Society* 4 (1): 2056305118763349. https://doi.org/10.1177/2056305118763349.

Bro, P., and F. Wallberg. 2014. "Digital Gatekeeping." *Digital Journalism* 2 (3): 446–454. https://doi.org/10.1080/21670811.2014.895507.

Chen, G. M., P. Pain, V. Y. Chen, M. Mekelburg, N. Springer, and F. Troger. 2020. "'You Really Have to Have a Thick Skin': A Cross-cultural Perspective on How Online Harassment Influences Female Journalists." *Journalism* 21 (7): 877–895. https://doi.org/10.1177/1464884918768500.

Cheruiyot, D. 2022. "Comparing Risks to Journalism: Media Criticism in the Digital Hate." *Digital Journalism* 12 (3): 1–20. https://doi.org/10.1080/21670811.2022.2030243.

Christin, A., M. S. Bernstein, J. T. Hancock, C. Jia, M. N. Mado, J. L. Tsai, and C. Xu. 2024. "Internal Fractures: The Competing Logics of Social Media Platforms." *Social Media + Society* 10 (3): 20563051241274668. https://doi.org/10.1177/20563051241274668.

Cinnamon, M. 2024. *Interventions for Online Harassment of Journalists*. Cambridge, Massachusets: Berkman Klein Center, Harvard University.

Collins, R.. 2014. " Interaction Ritual Chains and Collective Effervescence." In *Collective emotions: Perspectives from psychology, philosophy, and sociology*, edited by C. von Scheve and M. Salmela, 299–311. Oxford: Oxford University Press.

Davies, N. 2009. *Flat Earth News: An Award-winning Reporter Exposes Falsehood, Distortion and Propaganda in the Global Media*. Paperback-1 Jan 2009 ed. New York, NY: Vintage Books.

Dean, J. 2021. "Faces as Commons: The Secondary Visuality of Communicative Capitalism." In *The Aesthetics and Politics of the Online Self: A Savage Journey into the Heart of Digital Cultures*, edited by D. Della Ratta, G. Lovink, T. Numerico, and P. Sarram, 357–370. Cham: Palgrave MacMillan.

Dobber, T., and M. Hameleers. 2025. "The Social Media Comment Section as an Unruly Public Arena: How Comment Reading Erodes Trust in News Media." *Electronic News* 19 (1): 3–18. https://doi.org/10.1177/19312431241268011.

Downes, E. J., and S. J. McMillan. 2000. "Defining Interactivity: A Qualitative Identification of Key Dimensions." *New Media & Society* 2 (2): 157–179. https://doi.org/10.1177/14614440022225751.

Ganesh, B. 2018. "The Ungovernability of Digital Hate Culture." *Journal of International Affairs* 71 (2): 30–49.

Hanitzsch, T., A. Van Dalen, and N. Steindl. 2018. "Caught in the Nexus: A Comparative and Longitudinal Analysis of Public Trust in the Press." *The International Journal of Press/Politics* 23 (1): 3–23. https://doi.org/10.1177/1940161217740695.

Hayes, A. S. 2008. *Press Critics Are the Fifth Estate: Media Watchdogs in America*. London UK: Bloomsbury Academic.

Hendrickx, J., and M. Opgenhaffen. 2024. "Introduction: Understanding Social Media Journalism." *Journalism Studies* 25 (9): 919–930. https://doi.org/10.1080/1461670X.2024.2368861.

Honey, C., and S. C. Herring. 2009. "Beyond Microblogging: Conversation and Collaboration via Twitter." In *2009 42nd Hawaii International Conference on System Sciences*, 1–10. IEEE.

Janowitz, M. 1975. "Professional Models in Journalism: The Gatekeeper and the Advocate." *Journalism Quarterly* 52 (4): 618–626. https://doi.org/10.1177/107769907505200402.

Kalogeropoulos, A., R. Fletcher, and R. K. Nielsen. (2018) 2019. "News Brand Attribution in Distributed Environments: Do People Know Where They Get Their News?" *New Media & Society* 21 (3): 583–601. https://doi.org/10.1177/1461444818801313.

Kalogeropoulos, A., J. Suiter, L. Udris, and M. Eisenegger. 2019. "News Media Trust and News Consumption: Factors Related to Trust in News in 35 Countries." *International Journal of Communication* 13:22.

Leavitt, A. 2013. "From #FollowFriday to YOLO: Exploring the Cultural Salience of Twitter Memes." In *Twitter and Society*, edited by K. Weller, A. Bruns, J. Burgess, M. Mahrt, and C. Puschmann, 137–154. New York NY: Peter Lang Publishing.

Lewis, R., A. E. Marwick, and W. C. Partin. 2021. "'We Dissect Stupidity and Respond to It': Response Videos and Networked Harassment on YouTube." *American Behavioral Scientist* 65 (5): 735–756. https://doi.org/10.1177/0002764221989781.

Marwick, A. E. 2021. "Morally Motivated Networked Harassment as Normative Reinforcement." *Social Media + Society* 7 (2): 20563051211021378. https://doi.org/10.1177/20563051211021378.

Marwick, A., and d. boyd. 2011. "I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience." *New Media & Society* 13 (1): 114–133. https://doi.org/10.1177/1461444810365313.

Meeks, L. 2020. "Defining the Enemy: How Donald Trump Frames the News Media." *Journalism & Mass Communication Quarterly* 97 (1): 211–234. https://doi.org/10.1177/1077699019857676.

Miller, K. C. 2021. "Hostility toward the Press: A Synthesis of Terms, Research, and Future Directions in Examining Harassment of Journalists." *Digital Journalism* 11 (7): 1–20. https://doi.org/10.1080/21670811.2021.1991824.

Miller, K. C., and S. C. Lewis. 2022. "Journalists, Harassment, and Emotional Labor: The Case of Women in On-air Roles at US Local Television Stations." *Journalism* 23 (1): 79–97. https://doi.org/10.1177/1464884919899016.

Nieborg, D. B., and T. Poell. 2018. "The Platformization of Cultural Production: Theorizing the Contingent Cultural Commodity." *New Media & Society* 20 (11): 4275–4292. https://doi.org/10.1177/1461444818769694.

Ong, J. C., and R. Tapsell. 2022. "Demystifying Disinformation Shadow Economies: Fake News Work Models in Indonesia and the Philippines." *Asian Journal of Communication* 32 (3): 251–267. https://doi.org/10.1080/01292986.2021.1971270.

Park, S., C. Fisher, T. Flew, and U. Dulleck. 2020. "Global Mistrust in News: The Impact of Social Media on Trust." *International Journal on Media Management* 22 (2): 83–96. https://doi.org/10.1080/14241277.2020.1799794.

Parry, K. 2022. *A Theory of Assembly: From Museums to Memes*. Minneaopolis: University of Minnesota Press.

Peeters, S. 2021. "Digital Methods Initiative: Zeeschuimer." https://github.com/digitalmethodsinitiative/zeeschuimer.

Peeters, S., and S. Hagen. 2022. "The 4CAT Capture and Analysis Toolkit: A Modular Tool for Transparent and Traceable Social Media Research." *Computational Communication Research* 4 (2): 571–589. https://doi.org/10.5117/CCR2022.2.007.HAGE.

Petre, C. 2021. *All the News That's Fit to Click: How Metrics Are Transforming the Work of Journalists*. Princeton, New Jersey: Princeton University Press.

Phillips, W., and R. M. Milner. 2018. *The Ambivalent Internet: Mischief, Oddity, and Antagonism Online*. Hoboken, New Jersey: John Wiley & Sons.

Phillips, W., and R. M. Milner. 2021. *You Are Here: A Field Guide for Navigating Polarized Speech, Conspiracy Theories, and Our Polluted Media Landscape*. Cambridge, Massachusets: MIT Press.

Pintak, L., B. J. Bowe, and J. Albright. 2021. "Influencers, Amplifiers, and Icons: A Systematic Approach to Understanding the Roles of Islamophobic Actors on Twitter." *Journalism & Mass Communication Quarterly* 99 (4): 10776990211031568. https://doi.org/10.1177/10776990211031567.

Posetti, J., N. Shabbir, O. Douglas, and B. Gardiner. 2022. *The Chilling: A Global Study of Online Violence against Women Journalists*. Washington DC: International Center for Journalists (ICFJ) in partnership with UNESCO.

Rathje, S., J. J. Van Bavel, and S. van der Linden. 2021. "Out-group Animosity Drives Engagement on Social Media." *Proceedings of the National Academy of Sciences* 118 (26): e2024292118. https://doi.org/10.1073/pnas.2024292118.

Robertson, C. E., K. S. Del Rosario, and J. J. Van Bavel. 2024. "Inside the Funhouse Mirror Factory: How Social Media Distorts Perceptions of Norms." *Current Opinion in Psychology* 60:101918. https://doi.org/10.1016/j.copsyc.2024.101918.

Rogers, R. 2018. "Digital Traces in Context| Otherwise Engaged: Social Media from Vanity Metrics to Critical Analytics." *International Journal of Communication* 12:23.

Shepherd, T., A. Harvey, T. Jordan, S. Srauy, and K. Miltner. 2015. "Histories of Hating." *Social Media + Society* 1 (2): 2056305115603997. https://doi.org/10.1177/2056305115603997.

Shoemaker, P. 1991. *Gatekeeping*. Newbury Park, California: Sage Publications.

Shoemaker, P. J., T. P. Vos, and S. D. Reese. 2008. "Journalists as Gatekeepers." In *The Handbook of Journalism Studies*, edited by K. Wahl-Jorgensen, and T. Hanitzsch, 73–87. London: Routledge.

Siapera, E. 2019. "Organised and Ambient Digital Racism: Multidirectional Flows in the Irish Digital Sphere." *Open Library of Humanities* 5 (1): 13. https://doi.org/10.16995/olh.405.

Song, Y., Q. Lin, K. H. Kwon, C. H. Y. Choy, and R. Xu. 2022. "Contagion of Offensive Speech Online: An Interactional Analysis of Political Swearing." *Computers in Human Behavior* 127:107046. https://doi.org/10.1016/j.chb.2021.107046.

Strömbäck, J., and M. Åkerlund. 2025. "This Isn't Journalism, It's Propaganda! Patterns of News Media Bias Accusations on Twitter, 2010–2020." *Digital Journalism* 13:1–20. https://doi.org/10.1080/21670811.2024.2435574.

Törnberg, A., and P. Törnberg. 2024. "From Echo Chambers to Digital Campfires: The Making of an Online Community of Hate in Stormfront." In *Social Processes of Online Hate*, edited by J. Walther and R. Rice, 93–119. London: Routledge.

Van Dijck, J., T. Poell, and M. de. Waal, eds. 2018. *The Platform Society: Public Values in a Connective World*. Oxford: Oxford University Press.

Wahl-Jorgensen, K. 2019. *Emotions, Media and Politics*. 1st ed. Cambridge UK: Polity.

Waisbord, S. 2020a. "Mob Censorship: Online Harassment of US Journalists in Times of Digital Hate and Populism." *Digital Journalism* 8 (8): 1030–1046. https://doi.org/10.1080/21670811.2020.1818111.

Waisbord, S. 2020b. "Trolling Journalists and the Risks of Digital Publicity." *Journalism Practice* 16 (5): 1–17. https://doi.org/10.1080/17512786.2020.1827450.

Wischerath, D., R. Dielesen, D. Leenstra, B. Uilen, Z. Noorani, L. G. Amaya, M. Geboers, et al. 2022. Mapping Mobs: Technological Affordances, Metrics, and Digital Violence Against Journalists. https://www.digitalmethods.net/Dmi/mappingthemobs.

Zappavigna, M. 2015. "Searchable Talk: The Linguistic Functions of Hashtags." *Social Semiotics* 25 (3): 274–291. https://doi.org/10.1080/10350330.2014.996948.

Zhou, K., T. Wilson, K. Starbird, and E. S. Spiro. 2023. "Spotlight Tweets: A Lens for Exploring Attention Dynamics within Online Sensemaking during Crisis Events." *ACM Transactions on Social Computing* 6 (1–2) June 2023: 33, Article 2. https://doi.org/10.1145/3577213.