Global Media Journal
German Edition

BerlinUP Journals

Peer-Reviewed Original Article

Deciphering Digital Hate: Assessing the Evidence between Online Speech and Offline Violence in Africa

Nicole Stremlau, Caitlyn McGeer & Marlene Straub

Abstract: From Ethiopia to Sudan, there has been significant concern about the role of hate speech and incitement on social media to promote offline violence and, at its most extreme, genocide. These questions have become more urgent with the growth of large language models and Artificial Intelligence that are increasingly shaping online speech and may amplify existing concerns. In this paper, we interrogate the assumptions and myths about the causal link between online speech and its impact on the offline world by evaluating the empirical evidence. Overall, we found that there is limited evidence pointing to this direct association and, in line with broader literature on the underlying causes of violence, our review points to longer-term contextual, historical, and economic factors that often drive conflict, particularly in Africa. We conclude by identifying major evidence gaps and highlighting the need for caution when attributing the impact of online hate speech on violence.

Keywords: online speech, violence, hate speech, content moderation, Africa, social media, conflict

Author information:
Nicole Stremlau is Head of the Programme in Comparative Media Law and Policy, Centre for Socio-Legal Studies, University of Oxford and Research Professor, School of Communications, University of Johannesburg.
For more information: https://orcid.org/0000-0002-2728-2293
Email: nicole.stremlau@csls.ox.ac.uk

Caitlyn McGeer is a postdoctoral researcher at the University of Oxford. She focuses on the use of artificial intelligence and social media in conflict settings in Africa. Her research interests deal with gender and technology, human rights, and security. Caitlyn is currently working on the ERC-funded ConflictNet project within the Programme in Comparative Media Law and Policy at the Centre for Socio-Legal Studies.
Email: caitlyn.mcgeer@csls.ox.ac.uk

Global Media Journal – German Edition
Vol. 13(2) 2023
ISSN 2196-4807
DOI: 10.60678/gmj-de.v13i2.280

Marlene Straub was a Research Assistant on the ConflictNet project at the Centre for Socio-Legal Studies at the University of Oxford. She currently works for the German Federal Ministry of Digital Affairs and Transport on Foundational Issues of European Digital Policy.
Email: marlenesimranstraub@gmail.com

Acknowledgement:
This research is part of the ConflictNet project (The Politics and Practice of Social Media in Conflict). It has been funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 716686, ConflictNET).

To cite this article: Stremlau, Nicole; McGeer, Caitlyn & Straub, Marlene (2023). Deciphering Digital Hate: Assessing the Evidence between Online Speech and Offline Violence in Africa. *Global Media Journal – German Edition, 13*(2), DOI: https://doi.org/10.60678/gmj-de.v13i2.280

Introduction

The academic and policy discourse around the impact of social media on issues such as democracy, political polarization, and violent conflict has fundamentally shifted over the last several years. Much of the early literature on the internet's impact was positive, highlighting the opportunities for broadening participation (Loader et al., 2014), connecting communities (Castells, 2011), and evading more traditional forms of media censorship in illiberal regimes (Aouragh & Alexander, 2011). Platforms, such as Facebook, promoted their potentially peacemaking powers, and were described by their founders as tools for bringing communities together, for reducing divisions, and increasing understanding. It was just a few years ago that Facebook CEO Mark Zuckerberg was in Colombia to launch internet.org, an organization that has sought to bring free or affordable connectivity to those without internet access, when he lauded the potential for his platform to bring peace. In reference to the decades-long conflict with the FARC rebels, he argued that "a lot of conflicts are caused by misunderstandings", and that "the internet as a whole and social media will bring reconciliation and peace" (Phys.org, 2015).

With growing concerns about the role of social media in ongoing genocides and violence from Ethiopia to Myanmar to Ukraine, the debate has significantly changed. In June 2019, UN Secretary General Antonio Guterres warned against the stigma and discrimination that may be propagated on social media. He argued that "the internet and social media have turbocharged hate speech, enabling it to spread like wildfire across borders", and highlighted how "words can be weaponized and cause physical harm and the escalation from hate speech to violence has played a significant role in the most horrific and tragic crimes of the modern age, from the antisemitism driving the Holocaust, to the 1994 genocide against the Tutsi in Rwanda" (United Nations, 2019). But how much do we really know about the connection between online speech and offline violence? What is simply assumed and what do we have evidence for?

Concerns around the specific role of platforms such as Facebook in conflict-affected situations have accelerated, as more evidence emerges to suggest inequalities around online content moderation, particularly for smaller and poorer countries in the global south (De Gregorio & Stremlau, 2023). This is both in terms of resources devoted to human moderators working in local languages such as Amharic in Ethiopia or Burmese in Myanmar, as well as the absence, or near absence, of machine learning for such low-resource languages, meaning that the automated content moderation systems that do much of the work in the global north are ineffective (De Gregorio & Stremlau, 2023).

These concerns have been at the fore of a succession of reports about the failure of social media companies, and Facebook in particular, to effectively respond to allegations that the platform had a role in inciting violence during the genocidal

campaign against the Rohingya Muslim minority. The 2018 United Nations Independent International Fact-Finding Mission on Myanmar extensively outlined the types of hate speech that were common, mostly in the form of paid advertisements on social media. Facebook has since accepted that it played a role in spreading such hate speech (Human Rights Council, 2018). But despite the eventual expansion of efforts to address these concerns on the part of the company, including by employing more human content reviewers fluent in local languages, and establishing a team dedicated to the country and crisis (Su, 2018), reports of incitement and extreme speech online have continued. This suggests that the challenges of addressing such a complex problem, at scale, exceed current efforts (Global Witness, 2022).

Similarly, the widespread reports of crimes against humanity and ethnic violence in Ethiopia have been accompanied by concerns about the role of social media and the inability of platforms to respond effectively. Ethiopia was brought to the fore during the widely publicized Facebook leaks, an internal set of documents that were provided to the US Congress by whistleblower Frances Haugen. In her testimony to Congress, she argued that Facebook was "literally fanning ethnic violence" in Ethiopia because of the failure of the company to effectively moderate its platform outside of the US and Europe: "[W]hat we saw in Myanmar and are now seeing in Ethiopia are only the opening chapters of a story so terrifying, no one wants to read the end of it" (Akinwotu, 2021). In both Ethiopia and Myanmar, there are echoes of the genocide in Rwanda and the notorious role that Radio Mille Collines was seen by many to have had inciting violence and mobilizing for the massacres (Des Forges, 2002). Indeed, many have been drawing this comparison. Headlines include "Africa can prevent Ethiopia from going down Rwanda's path" (Kissi, 2021) or "Will Ethiopia's genocide be worse than Rwanda's?" (Rubin, 2021), with the articles implicating hate speech as a main driver of the violence.

As the shift towards more critical views on the potentially destructive aspects of hate speech on social media continues to accelerate in public debate, there is an urgent need to interrogate the assumptions that underscore many of these headlines and claims. Causation often seems to be assumed - hate speech on social media causes a new episode, or increase in, violence - with little evidence offered. Or in other cases correlation appears to be conflated with causation because more people are on social media in certain areas and being exposed to hateful or polarizing content therefore it is a distinctive driver in a conflict, again often with little evidence offered. We agree that social media matters during violent conflict or in very different economic, social, and political contexts but it is not always obvious how, why or in what ways.

This article adds nuance to this debate by exploring and evaluating existing literature and evidence around social media's impact, and the frequent claims of almost immediate impact, on inflaming violence, particularly in countries in Africa. Doing so forces reflection on not only what we know about conflict, but how social media effects are understood, or theorized, in conflict-affected situations. It also calls for critically reflecting on the general portrayal of social media and conflict in Africa, in

comparison with other regions of the world. Our focus on Africa, and sub-Saharan Africa in particular, builds out of the extensive empirical research we have been conducting as part of a large European Research Council project on *The Politics and Practice of Social Media in Conflict in Africa*.

We used a semi-structured approach to identify and evaluate the evidence including a quasi-systematic literature review and reaching out to experts. Our focus was on information communications technologies (ICTs) which enable internet access (e.g. mobile phones) and internet-based applications (e.g. social media platforms). These are frequently intertwined as it is often difficult to disentangle mobile phone use from internet use (Stork et al., 2013) and network providers are oftentimes the same (Macías-Medellín & Atuesta, 2021).

Our findings identified a pool of texts that specifically examine the impact of ICTs on violence in sub-Saharan Africa. Overall, we found that there is limited evidence pointing to this direct association and, in line with broader literature on the underlying causes of violence, our review points to longer-term contextual, historical, and economic factors that often drive conflict. We are very much aware that this is part of the challenge of this field of research: technologies impact, and interact with pre-existing circumstances in a given place – e.g. political disenfranchisement or resource scarcity. Nevertheless, beyond our general conclusions on the strength of the body of evidence, several trends emerge in the literature that we explore in turn including (i) evidence around the reproduction of societal divisions and exclusions through network coverage, manifesting in a greater incidence of violence, (ii) infrastructural preconditions of ICTs that are seen to contribute to violent conflict; and (iii) geo-spatial or transnational dimensions of online speech and conflicts that are often overlooked in policy discussions.

Before discussing these themes within the literature, we situate the current discussion around social media and new technologies within the broader historical debates about media and conflict, particularly in Africa (Stremlau, 2018). This has, in many respects, been the precursor to how social media and conflict is approached.

From Media to Social Media and Conflict

Concern about the role of media, or in our case social media, in genocide has been longstanding.[1] Much of this is, by now, well known but several examples stand out and have shaped the common contemporary narrative and perception of communications in mass violence. During the first world war, and the Armenian genocide (1915-1923), for example, newspapers were actively leveraged by the Ottoman government to spread anti-Armenian propaganda, depicting Armenians as a threat to state stability because of their alleged collaboration with Russia or as infidels whose

[1] For a review of evidence around the role of media in conflict see Shoemaker & Stremlau (2014).

faith put them at odds with the majority of the population. Reporting on and coverage of mass deportations and systematic killings was largely suppressed (Akçam, 2012). Similarly, during the Holocaust (1933-1945) Nazi propaganda was actively espousing Nazi ideology and repeated narratives dehumanizing Jews, people with disabilities, and other targeted groups. Newspapers, radio, arts, and the film industry were leveraged to portray these communities as undermining the quest for the perfect Aryan society (Herf, 2006).

More recently, the 1990s conflicts in the Balkans and Rwanda re-ignited the debate over the association between media and conflict and the urgency to respond. In the Balkans (1992-1995), media, and radio in particular, was seen to have had a central role. In Bosnia and Herzegovina, Serbian media outlets portrayed Bosnian Muslims as a threat and were actively supporting genocidal narratives in Srebrenica and other regions (Thompson, 1999). It was during this time that the concept of information intervention came into its own whereby the international community (in the case of Bosnia, NATO) had an active role in bombing transmitters to forcibly take hateful propaganda and incitement to violence off the air (Metzl, 1997; Thompson & Price, 2003). Information intervention began to attempt to directly address the tension between the responsibility to protect, or the UN's principles on the prevention of genocide, and what is often an overwhelmingly US-led deference to freedom of expression. While the doctrine of information intervention failed to gain significant traction, it highlighted the perceived failures of when international actors do not act on media that might be involved with extreme violence (De Gregorio & Stremlau, 2021). It was in the aftermath of the Rwandan genocide in 1994, many assigned great responsibility to Rwandan media – principally the government-controlled Radio-Télévision Libre des Mille Collines (RTLM) – in inciting violence that escalated into the genocide of the Tutsi minority. Samantha Power famously claimed that "[k]illers often carried a machete in one hand and a transistor radio in the other" (2001). Indeed, all stakeholders – from international law scholars to Hollywood film producers – feature the radio as *the* symbol and paradigmatic cause of the genocide (Straus, 2007). Despite these strong claims, the role of radio in the Rwandan genocide also points to challenges of drawing direct connections between online hate speech or incitement and offline violence in places such as Myanmar and Ethiopia. While research and commentary in the aftermath of the genocide tended to focus on the perceived role of media, as time passed and more empirically grounded work emerged, the exact influence of RTLM in the genocide has been contested (Grzyb, 2019).

As Straus (2007) pointed out, the methods of empirical social science research – essentially, what this article is looking for – have scarcely been applied in the vast literature on the Rwandan genocide, "despite the presence of often quite strong claims about media effects [which] often assert or imply undifferentiated, direct, and massive media effects – effects that, if true, would be at odds with decades of political communications empirical research" (p. 610). We will return to this point of media effects in the context of social media in the conclusion.

Studies on the Rwandan genocide that have a more nuanced approach to media have argued, for example, that radio was an extension of state propaganda that was spread throughout all public institutions (McDoom, 2012), and that in this context, radio only had a "conditional and marginal effect" on a minority of perpetrators, who were already hard-liners before (Straus, 2007). Yanagizawa-Drott (2014) statistically explored the impact of radio coverage on the incidence of violence in the Rwandan genocide and concludes that approximately "10% of the total participation in the genocide, or approximately 51,000 prosecuted persons, was caused by the radio station" (p. 1950). While this modest percentage is in line with more critical analyses of RTLM's actual impact and is one of the rare studies to find a statistical connection between broadcasts and violence, Yanagizawa-Drott's (2014) approach is not without criticism. Using estimates of media coverage as a proxy for consumption, the actual variable of interest, has been proven to be flawed by several studies which qualitatively investigate actual consumption patterns on the ground (Danning, 2018). Relative to radio exposure, face-to-face interactions with peers and social ties as well as horizontal pressure in the community have consistently been found to be more catalysing for participation in and support of violence (Fujii, 2011; McDoom, 2013; Straus, 2007). While it would be wrong to say that radio played no role at all, rumours, some of which derived from broadcasts, appeared to be more powerful mobilizing forces (Li, 2004; Mironko, 2007).

Determining how to disentangle these forces such as radio broadcasts from rumour and coverage from actual consumption, is difficult (Danning, 2018). Because of these research challenges, assumptions which are not necessarily underpinned empirically spread in scholarly and popular discourse creating myths of causality and potentially undermining research in the future. If RTLM was ascribed lethal influence, the international community's refusal to employ information interventions or radio jamming certainly was a failure in preventing the genocide, as it might have had an impact in mitigating the catastrophic events that occurred (Metzl, 1997). Indeed, the causal link between either radio in the conflict in Rwanda, or social media in Ethiopia or Myanmar, and the ensuing violence is not entirely clear.

Nevertheless, the view that social media has been a central, if not the central factor, has been reinforced by recent lawsuits. Rohingya from the diaspora, have launched cases in the US and UK, seeking damages from Facebook for allowing incitement to violence to flourish on its platform (Economist, 2021) and there has been a 1.6 billion USD lawsuit in Kenya, which hosts Facebook's East African content moderation hub, by several plaintiffs from Ethiopia. One of the plaintiffs, Abraham Meareg Amare, argues that Facebook's algorithm and failures to effectively moderate content are directly responsible for his father's death (Olson, 2022).

A UN fact-finding mission found Facebook to be a "significant" instrument for those wishing to incite hatred against the Rohingya minority in Myanmar, and ascribed a "determining role" to the platform (Human Rights Council, 2018). The report

Facebook commissioned went even further, concluding that hateful propaganda spread on the platform had been linked to offline violence but also noted that "the actual relationship between content posted on Facebook and offline harm is not fully understood" (BSR, 2018). In this context, the focus lies on Facebook, as "for most users, Facebook is the Internet" (Human Rights Council, 2018).

Contrary to some of the challenges faced by the research discussed in this article, Facebook has access to all kinds of data on how content was interacted with on its platform. The company is actively fighting to keep such data off the record including in the proceedings before the International Court of Justice, which is investigating the genocidal crimes committed against the Rohingya people (Towey, 2021). Such data could be critical in deepening our understanding of the association between online speech and offline violence.

Conceptual and Methodological Challenges in the Literature

We now turn towards examining empirical research that has sought to explore the link between online speech and offline violence, particularly in cases of extreme conflict in Africa. In developing and refining our methodology, we drew on the UK Department for International Development (DFID)'s guidance on 'Assessing the Strength of Evidence' (2014). We pursued a multi-pronged semi-systematic approach. We began by adopting a quasi-systematic approach to identify literature using key search terms in scholarly database searches to gather initial evidence (Higgins et al., 2019, p. 4). From there, we conducted snowball citation searches among the sample texts identified for inclusion. This latter approach enabled us to expand beyond the traditional confines a systemic review to ensure we captured as much literature as possible. We also conducted an expert survey of more than 50 experts working in the field of media, or ICTs, and conflict including academics and practitioners for international organizations.

We sought relevant literature from multiple scholarly databases, including Google Scholar, Scopus, Crossref and Web of Science. We used a variety and combinations of key terms including, for example: 'Africa', along with countries with recent occurrences of violence (i.e. 'Kenya', 'Uganda', 'Ethiopia'), 'violence', 'violent action', 'civil unrest', 'crime', 'conflict', 'mobilisation/mobilization', 'hate', 'online harm', 'offline harm', 'racism', 'ethnic', 'hate', 'mobile phone', 'ICT', 'technology', 'social network', 'social media', 'messenger', 'coverage', 'internet'. The database searches yielded a total of n=25,621 pieces of literature. With a Python script, n=11,985 duplicates and erroneous entries were removed, leaving n=13,636 unique pieces of literature.

The link between online and offline harms has been studied by a variety of disciplines, from information science to psychology, reflecting varied methodologies. Based on DFID's (2014) approach, however, we filtered to only include literature that involved empirical research, be it qualitative or quantitative. After a preliminary manual review of the remaining texts, based on titles and abstracts, n=489 publications were left. Through a closer reading, requirements for inclusion were applied. Further guided by DFID's (2014) approach, we focused on research that appeared in a peer-reviewed publication or in form of so-called 'grey literature', produced outside traditional academic publishing. The pieces had to address the 'harm' concerned in large-scale violence or conflict (as opposed to say, isolated incidences of violent extremism or smaller scale violent protests). Our geographic filter included African states or pieces that addressed Africa as a whole. Publications addressing other countries or continents were excluded, due to the potential for unique ICT structures, inferred from mobile phone penetration and other contextual factors. In the end, n=11 key publications remained, which were published between 2013 and 2021[2]. This period corresponds with one of rapid ICT diffusion in Africa.

This low number of articles (n=11) from the database review was surprising, so it was validated again after reviewing our filtering process. There were several recurring reasons why literature was excluded from the scope of the review. This review is concerned with violent conflict offline, as can occur, for example, in an uprising against a government force or genocide. A considerable proportion of excluded literature on collective action, however, did not address violence, only mobilization and potential pathways towards extremist radicalization. We excluded a significant amount of literature on protests in Africa – many of which were violent or had an element of violence – as violence was only mentioned in passing, if at all. If texts did not address *violent* collective action specifically, only the impacts of ICTs on collective action more generally, they were also excluded. Inversely, the large body of literature which considered how ICTs could be used to eliminate violence and promote democracy, to manage or prevent conflicts, was significant and also eliminated (e. g. Trujillo et al., 2014). Many more texts were removed for not involving empirical research. These largely focused on the normative framing, giving a historical account (e.g. Mutahi & Kimari, 2017) or being descriptive otherwise (e. g. Dafoe & Lyall, 2015; Gohdes, 2018).[3]

[2] The texts that were filtered included Pierskalla & Hollenbach (2013), Ingrams (2015), Bailard (2015), Weidmann (2015), Camber Warren (2015), Hollenbach & Pierskalla (2017), Absher & Grier (2019), Njuguna, Gikandi, Kathuri-Ogola & Kabaria-Muriithi (2020), Manacorda & Tesei (2020), Ackermann, Churchill & Smyth (2021), Macías-Medellín & Atuesta (2021).

[3] There are limitations to our methodology, of course; some of which resulted in the filtering between the initial and inclusion sample. Primarily, there is the issue of the search terms. We consciously included a greater number of search terms, to minimize the risk of overlooking relevant research, at the risk of generating a considerably over-reaching initial sample. It is also possible that, throughout the review process of this large sample, relevant articles were falsely excluded. Furthermore, there is an inherent exclusionary effect in using exclusively English-language search terms, which primarily generated English-language results and may have resulted in the omission of relevant non-English literature. Findings from Francophone Africa, for example, are not captured in this review, which is limiting.

To ensure we did not miss any relevant publications in our review process, we engaged in snowball sampling from the n=11 articles for inclusion. We collated the n=667 citations across the sample of texts and manually cross-referenced this spreadsheet with the original findings from the review (n=13,636 after filtering out duplicates and errors). Through this process, no additional literature that met the inclusion criteria was added that was not included in the original sample. Considering the small size of the original sample, this was reassuring.

To extend the reach of our review we triangulated our findings with experts. We contacted n=69 experts in the field (academic and practitioners) in early 2023 to ask for their recommendations for up to n=5 pieces of literature relevant to the review. This process garnered n=85 recommended pieces of literature after removing duplicates. There were only three publications that were recommended by multiple experts. One was a discussion paper, *The Chilling*, published by UNESCO (Posetti et al.) in 2021 on global trends in online violence against women journalists, which was not directly relevant for our review.[4] Another was an article outside of our geographic scope by Müller and Schwarz (2020), *Fanning the Flames of Hate: Social Media and Hate Crime*, that focused on how online hate speech directed at refugees in Germany relates to offline violence against this community. Finally, Pierskalla and Hollenbach's (2013) article *Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa* was recommended by multiple experts, but it was already included in the review from the database searches. Notably, Bailard's (2015) article*, Ethnic Conflict Goes Mobile: Mobile Technology's Effect on the Opportunities and Motivations for Violent Collective Actio*n and Ackermann et al.'s (2021) article, *Mobile Phone Coverage and Violent Conflict*, both while only recommended by one expert, were also already included in the review.

Beyond these, the texts recommended ranged quite drastically in focus and geographical scope, from incel violence to the making of modern Ethiopia between 1896-1974. Of the expert recommendations, n=21 were further reviewed and two met the criteria for inclusion including Pelican et al.'s (2022) working paper on extreme speech and violence in Anglophone Cameroon and Fokou et al.'s (2022) article on online disinformation and offline xenophobic violence in South Africa.

Trends in the Texts: Geography, Temporality, Methodology, and Underlying Assumptions

As exemplified by the low number of texts generated by the sampling process, there is an apparent dearth of empirical research probing the connection between online harm and offline violence. Of those included in the review, there were also important gaps within these texts that limit the evidential value of this review. For instance,

[4] While the report featured findings from a survey with women from 17 countries and 3 of which were in our geographic focus (Kenya, Nigeria, and South Africa), the findings were largely presented as a whole, making it difficult to disentangle the experience of African journalists from others.

the authors extensively draw on literature studying ICTs' influence on collective action and protest mobilisation, on violence reduction or prevention, and their capacity to strengthen democracies. They contextualise their research on ICTs' possible effects on violence within their potential for economic development or improving the access to education. Most notably, the authors heavily cross-reference. The Pierskalla and Hollenbach (2013) study is cited by most of the other sample texts, and most texts reference at least one or two other sample texts.

Equally important, only a handful of studies differentiate between internet-enabled and non-internet-enabled mobile devices (e.g. Ingrams, 2015; Macías-Medellín & Atuesta, 2021; Warren, 2015). This is problematic, considering that the studies which do differentiate find distinctly different effects across both categories. Ingrams (2015), for example, finds that internet-enabled mobile phones may facilitate civic deviance, including violent conflict, whereas non-internet-enabled mobile phones have the opposite effect. The differentiation is particularly relevant in many African countries, where approximately twice as many people have mobile cellular subscriptions than use the internet, an even smaller subset of whom use social media (International Telecommunication Union, 2019; Kemp, 2021). There is ample evidence that network providers who offer both telecommunications and internet coverage are the same, and that the infrastructure for the latter has outpaced the former in many parts of Africa. Even the earliest study in scope, Pierskalla and Hollenbach (2013), researches a period from 2007 to 2009 when 2G technology, which allows for basic internet access, had been introduced in Africa (Manacorda & Tesei, 2020). Certainly, the ability to disaggregate technologies will be limited by what data is available to researchers. However, an acknowledgement of the potential for different effects or an effort to mitigate a conflation thereof is needed. In many studies, no mention was made of this, and many texts exploring mobile phones' overall effect will have unintentionally, at least partially, studied the internet's effect, and vice versa. Without differentiation, there is no way to unravel distinct effects.

Several studies also relied on network coverage data, at more or less granular geographic resolutions. This data, as mentioned, is generally provided by industry bodies, who might have an interest in portraying their members' network coverage favourably, and whose data collection process is not usually transparent. Coverage does also not reflect actual usage of mobile devices. While coverage may reflect demand, to be able to isolate the effect of online harms on offline violence, more granular data is necessary to draw conclusions on, for example, an increase in network use around an uptick in violent events, perhaps even on the domains visited. Companies like Facebook may be able to offer such data, but refuse (Towey, 2021). Ingrams (2015) further notes a crucial limitation in the lack of detail on the variability of mobile phone uses and usage, treated as single construct. We do not know *how* mobile phones are used, and indeed, as we have learnt in the years since Rwanda, the rumour mill is likely more powerful than the technological source of information – coverage or usage data does not capture this.

Similarly, most texts in scope cover relatively short periods of time. Pierskalla and Hollenbach (2013), for example, cover the years 2007 to 2009 and base their findings on a snapshot, even though their datasets cover longer timeframes. Ackerman et al. (2021) are one of the few exceptions in considering an 11-year time-period. On balance, if not in the case of natural experiments (i.e., Absher & Grier, 2019), research which explores longitudinal developments are desirable. Especially in the case of statistical modelling, the simplistic assumption is often that effects manifest with a lag of one temporal unit and are consistent. In contrast, as Macías-Medellín and Atuesta (2021) show, different ICTs have varying effects at the onset of use versus once they are established. To extrapolate general statements about the effect of the internet on offline violence at the pan-African meta-level over short timeframes is not very informative, as different African states are at vastly different stages of infrastructural developments, particularly at present. More detailed data on the use of individual ICT services or devices would be very beneficial, to tie developments over time to specific use cases. Importantly, also, the literature in scope pays less attention to more recent periods. Potentially, this may be due to the messiness of disentangling differences in internet-enabled versus non-enabled mobile phones, but there is a real need to delve into more recent uses of the internet, and potential impacts on offline violence, precisely because of the overlapping messiness and increasing use cases, and the rapid development of the internet.

It is also important to note that conversations on how the internet or new technologies exacerbates online harms generally refer to the role of social media platforms. Except for one study, the literature turned up in our review not directly address these but focuses on broad mobile phone or internet coverage, not individuals' use thereof, and explores developments over time periods several years ago, when social media platforms did not look as they do now. In effect, it appears as if the body of literature on the connection between online and offline harms does not address what is commonly imagined but this also may reflect the time it takes for new studies, for example on Ethiopia or Myanmar, to be peer-reviewed and published.

There are also limitations in data sets the literature used. The texts in scope predominantly drew on large datasets generated by external researchers, covering (i) the incidence of violent conflict over time, (ii) network coverage over time, and (iii) geo-spatial data to divide regions for study. Many of the respective datasets were used by more than one study. On the incidence of conflict, Uppsala University's Department of Peace and Conflict Research has collected data on global violent conflict for the past 40 years. Its Uppsala Conflict Data Program (UCDP) dataset (Pettersson et al., 2021) was used in the majority of the texts. Data on network coverage was primarily provided by industry bodies, such as the Global System for Mobile Communications (GSM) Association, which offered Manacorda and Tesei (2020) data on the availability of signal between 1998 and 2012 at a geographically precise level. To be able to map the events data with coverage data, most studies relied on so-called 'grid cells', dividing the world into geographical squares. The PRIO-GRID

dataset was the most popular, used by half of the studies.[5] There was little variation between these studies in research design, each drawing on a dataset reflective of each of these categories.

Notably, the most significant study departing from this norm, was highly localized and explored electoral violence in Mathare Constituency, Nairobi County, Kenya, for which it generated its own dataset on a basis of questionnaires, interviews and focus group discussion (Njuguna et al., 2020). This, coincidentally, was also the only text in scope written by African authors and published in an African journal. A central criticism of the scholarship which emerged post-Rwanda, -Darfur and -Myanmar was the over-representation of 'outsider' Western scholars, whose perspective was very removed from local contexts, causing a variety of problems (Ibreck & de Waal, 2022). At the very least, the pattern of Western over-representation seems to have been reproduced in the sample, whereas the reliance on largely anonymous datasets at a higher level seems to support the historical pattern whereby grant theories are carelessly proposed "could only ever make sense from a distance" (Ibreck & de Waal, 2022, p. 84). It is likewise unsurprising that the single study which draws data from conversations with people affected by online and offline violence in Africa was conducted by researchers embedded in these contexts.

Furthermore, when it came to methods, the majority of the studies were quantitative, reflecting broader trends in how evidence reviews define evidence. For one, most datasets on the incidence of violent events rely on sources such as news reports for evidence of a violent event occurring. This measure comes with many limitations. Weidmann (2016) dedicated a paper to examining just one: "[i]f cellphone coverage makes the reporting of violence more likely, this can lead to measurement error in the dependent variable that is correlated with the main independent variable [- cellphone coverage]." (p.210) Without wholly refuting Pierskalla and Hollenbach's (2013) study, Weidmann (2016) concluded that their results may have been skewed by reporting bias. Bailard (2015) may have encountered a similar problem: violence in rural regions may have been especially underreported before coverage extended, therefore, the finding that greater coverage in rural areas led to greater violence may be a symptom of reporting bias. Warren (2015) controls for this potential by relying on a secondary dataset which mitigates the risk of reporting bias by extensively using secondary sources. These carry their own risk, emanating from the sources' language. If data was recorded in English, the "discourse in the local language may follow very different dynamics [...] and different languages are likely to reflect entirely

[5] The reliance on geographic grid-cells comes with an inevitable statistical limitation. Grid-cells are pre-specified in datasets such as PRIO-GRID, raising the question whether empirical inferences would have differed, had the grid been constructed differently or shifted. Among geographers, this issue is called the 'modifiable areal unit problem' and well-known for its ability to "generate severe statistical biases, with widely varying directions and intensities, rendering multivariate spatial regression estimates inherently unreliable" (Warren, 2015, p. 299; see also Openshaw, 1984). Most studies in question use statistical measures which are more complex than simple univariate or bivariate regressions. In such contexts, the "modifiable areal unit problem is shown to be essentially unpredictable in its intensity and effects" (Fotheringham & Wong, 1991, p. 1025).

different policy positions and viewpoints" (Gohdes, 2018, p. 99). While this limitation is known, little can be done to investigate it. Even publicly available datasets, such as the Uppsala Conflict Data Program (UCDP,), offer limited insight into the underlying reports which informed the data. In effect, reporting bias in events data on the incidence of violent events suffers from dual problems: (i) reporting bias interfering in the accuracy of research findings, and (ii) the accuracy or bias in the evidence which informed the events data generation.

Several assumptions are frequently made in the texts, which appear to ground the overall approaches to research. Almost every text in the sample builds its argument on literature from protest studies, as to how internet technologies can help individuals overcome two dual problems of collective action, arguing that access to the internet via mobile phones would (i) enable greater dissemination of information, and (ii) enhance mobilization by reducing collective action and coordination problems. An assumption inherent to this is that all violent conflict stems from the same mass mobilisation logic that underpins collective action. This need not necessarily be the case. Even if this fallacy was ignored, none of the texts justify why the framework from collective action should map directly onto *violent* collective action. Presumably, the thresholds differ between the two, as do organisational structures towards the differing ends.

Further research is also needed on whether the premises of improved dissemination of information hold. While it is plausible that greater connectivity helps eliminate information asymmetries, it is equally plausible that new ones are created, or old ones exacerbated. The internet is, after all, full of information warfare, propaganda, dis- and misinformation. Granted, they may lead to the same outcome – the eruption of violence – but the nature of the violence changes by which angle of the information mechanism individuals arrive at this outcome from. As mentioned, the Media Manipulation Casebook project consider the Ethiopian online discourse a 'war over the narrative' (Wilmot et al., 2021), in which all stakeholders produce and share their own information operations.

A central issue with the assumed logics of information diffusion and the reduction of collective action and coordination problems is that they are themselves rather broad and unspecific. Broadly speaking, the cited literature – Njuguna et al. (2020) being the exception – share similar theoretical frameworks and approaches. Yet the overarching mechanisms of information diffusion and coordination are so amorphous that it is quite difficult to extrapolate, as Dafoe and Lyall (2015) put it, "theory that is elaborate (has many testable implications), explicit, precise, logically developed, informed by intuition from extensive field experience, and grounded in the body of empirical findings" (p. 402). Evidently, granularity is missing in the data and theory to facilitate multiple testable mechanisms. The authors of the studies in scope are largely transparent about this shortcoming. Pierskalla and Hollenbach (2013) state that their "results only imply an association at the aggregate level of the spatial unit and do not reveal the exact causal mechanism" (p. 221), whereas Absher

and Grier (2019) caveat that they are only "estimating the net effect of all relevant mechanisms" (p. 6). Overall, this body of literature concludes that the spread of mobile phone use and internet access are manifesting in violent offline harms, which we explore next.

Connecting the unconnected: are ICTs changing conflict in Africa?

The included literature maps the progression in the field over almost a decade. While the earliest text begins recognizing that little is known about whether mobile phone technology has increased violent forms of collective action (Pierskalla & Hollenbach, 2013), the most recent texts continue this spirit. According to Ackermann et al. (2021), "[t]he evidence on the relationship between phones and violent conflict is mixed" (p. 270). This lasting ambiguity points towards overarching difficulties in studying the connection between the internet, online and offline violence.

While every article shared the premise that ICTs exert great and increasing influence on African societies in a of myriad ways, it was often framed within the context of the potential benefits ICTs can bring for development: in economic terms and for the financial sector, for education and empowerment purposes, the dissemination of health and other information through public institutions, the improvement of living standards more generally and so forth, reflecting the prevailing techno-solutionist or ICT for development in Africa. Most commonly, ICTs are ascribed considerable political influence in facilitating the dissemination of political information and encouraging democratic participation, but also as catalysts for collective action and civil unrest. Most texts reference ICTs ability to enhance mobilization by reducing collective action and coordination problems, drawing extensively on protest literature.

Mobile phone technology has widely been linked to facilitating collective action, due to its ability to help overcome various collective action and coordination problems (e.g. the free-riding problem). This factor is frequently mentioned as *the* underlying mechanism linking online to offline events in literature on mobilisation for collective action (Gohdes, 2018) reflecting the influence of the Arab Spring on this field. Two central mechanisms drawn from the body of evidence on mobilisation for collective action – (i) improved information dissemination and (ii) coordination – are referenced by most texts in the sample (i.e. Pierskalla & Hollenbach, 2013; Warren, 2015). There is a shared view throughout the texts that with the improved dissemination of information via horizontal, peer-to-peer channels, information asymmetries can be overcome, individuals can be empowered, and government propaganda can be undermined. Coordination can be improved as internet-enabled devices become cheaper and more accessible, easing communication across time and distant regions, by tightening group networks, raising in-group trust, cohesion, and by providing ways to monitor peers. Within our sample, the differences lie primarily in how these two mechanisms are methodologically approached.

Three key findings emerge: (i) the reproduction of societal divisions and exclusions through network coverage, particularly urban-rural, manifesting in a greater incidence of violence, (ii) infrastructural preconditions of ICTs that are seen to contribute to violent conflict; and (iii) geo-spatial or transnational dimensions of online speech and conflicts that are often overlooked, particularly in policy discussions.

The Urban-Rural Divide and the Reproduction of Societal Exclusions

The majority of texts in the sample acknowledge that access to network coverage and mobile technologies mirror distributions of power in society. At the most basic, it is commonly accepted that there is substantial variation in the geographic spread of internet access both between and within states (Warren, 2015). Pierskalla and Hollenbach (2013) find a significant effect of mobile phone coverage on violent conflict erupting, which is especially pronounced in areas with structural conditions which favour violence, such as a legacy of conflict. This effect is compounded by whether an ethnic group in an area faces political exclusion, even when controlling for these groups' potential undersupply with communication technologies and infrastructure. Put differently, through the spread of mobile phone technologies, ethnic groups which are excluded from political say are somehow empowered differently than the general population to engage in collective violent action. It is not clear what mechanisms may cause these different degrees of empowerment.

Bailard (2015, p. 334) expands on Pierskalla and Hollenbach's findings and maps the percentage of an ethnic group's territory covered by mobile signal, weighed by population density, on the probability of violent conflict erupting. She finds that "increased mobile signal coverage significantly increases the probability that a group will engage in conflict with its government in a given year" (Bailard, 2015, p. 334). An ethnic group's population and population density condition the impact, with smaller or intermediate-sized and the more sparsely distributed groups benefitting more from the introduction of mobile phones for mobilisation purposes. Bailard (2015), parallel to Pierskalla and Hollenbach (2013), finds that exclusion from political power, potentially tied to group size, also increases the likelihood of organising for collective violence. Likewise, but more broadly, Weidmann (2015) argues that the proportion of the population excluded from political power is positively related to the likelihood of conflict erupting, richer countries and democracies have lower risks of conflict, whereas large, potentially diverse countries are more at risk.

Manacorda and Tesei (2020) made an economic argument and found a statistically significant effect of ICT coverage on collective action during periods of recession or economic hardship. This effect was especially pronounced in areas with a legacy of past violence relative to comparatively peaceful areas, indicating that "citizens in high conflict regions may be more responsive to [economic] downturns" (Manacorda & Tesei, 2020, p. 560). A legacy of past conflict likely indicates a degree of socio-economic and political instability offering more fertile ground for disruptive new technologies to have a role in violence, if triggered by destabilising economic

challenges. Ackerman et al. (2021) have different findings and suggest that while economic growth dampens the effects of mobile phone coverage on conflict, the effect of coverage on violence erupting is positive, dominant, and nevertheless outweighs the effects of economic growth. In other words, mobile phone coverage contributes to violence, even when the economy is doing well. The literature does not address the role of government in violence (either as an actor or responder), only the circumstances in which groups might engage in violence.

Infrastructural Preconditions Contributing to Violence

Sub-Saharan Africa is characterised by infrastructures that differ among its countries and regions, as well as internationally. The order and timing of the introduction of different technologies has also varied greatly across the continent, with internet coverage outpacing 'old' technologies (Warren, 2015). Notably, throughout the investigations of the violence against the Rohingya minority in Myanmar, a central conclusion was that the exponential spread of the internet in an otherwise infrastructurally less developed country was a key factor in why the technology had a determining effect on violence (BSR, 2018, p. 24). An exponential spread in internet penetration via mobile phones met a population with low media literacy, in a context previously characterized *inter alia* by little free expression, in part, due to the government controlling the telecommunications sector. Such considerations are also found in the ways in which social media in the conflict in Ethiopia is discussed.

Within the literature we examined, Warren (2015) specifically controlled for infrastructural legacies and finds that "geographic regions with greater levels of penetration by cellular infrastructure experience systematically higher rates of collective violence" (p. 306), although these effects are greatly conditioned by prior levels of radio penetration, "with the greatest levels of violence observed in regions where the arrival of cellular technologies has occurred in the absence of prior penetration by radio infrastructure" (p. 306). Interestingly, Warren (2015) finds that this effect holds, even when controlling for factors such as regional wealth or societal inclusion of an ethnic group. This finding runs counter to the conclusions in the previous section, showing how complicated the picture is. Bailard (2015) similarly found that increased mobile signal coverage significantly raises the probability of violence erupting, which is found to be especially pronounced in areas with weak prior landline coverage. The effect, therefore, holds across both radio and landline penetration.

Considering how straightforward and significant Warren (2015) and Bailard's (2015) findings are, it is surprising that so few studies specifically control for infrastructural preconditions. Especially since, as the previous section showed, power imbalances in society determine how infrastructures will be distributed among its members. Indeed, Weidmann et al. (2016) find "a strong and persistent political bias in the allocation of Internet coverage across ethnic groups worldwide [...] which cannot be explained by economic or geographic factors" (p. 1151). They suggest that

internet allocation may be a service only exchanged to those politically favoured by the government, thus actively working against ICTs potential to be "liberation technologies".

The implications of this research, when attempting to identify what, for example, it might mean for a country like Ethiopia which is currently experiencing violence many are associating with social media are not entirely clear. Fixed landline and broadband subscriptions plummeted in recent years as mobile cellular subscriptions have grown exponentially but much of this has been in urban areas. The conclusion that internet infrastructure rapidly outpaced traditional telecommunications holds in the context of Ethiopia, for example. But more granular data at the regional level suggests that the sheer subscription numbers show that most regions which would require greater ICT coverage would receive it without prior penetration of other telecommunications infrastructure. Therefore, following the arguments of Warren (2015) and Baillard (2015) might suggest that the lack of communication infrastructure prior to cellular infrastructure leads to higher rates of violence, particularly in areas that have not had radio (Warren, 2015) or landline infrastructure (Bailard, 2015) and it might be assumed that some of the more recent violent conflict might be partially caused by a rapid rise in availability of peer-to-peer information, met by a population with comparatively low, or misunderstood, levels of media literacy. At the same time, however, we know that there has been an internet shutdown across much of the most heavily-affected conflict regions in Ethiopia, that, in some places has lasted for over two years. As our previous section discussed, context very much matters, and most observers of Ethiopian politics would situate the current violence within a series of unresolved conflicts stretching back more than a century to the formation of the Ethiopian state.

Spatial and Transnational Dimensions

At basic level, there is a spatial dimension to the effects of ICTs on violent conflict. Violence may spill over from a city to neighbouring villages, and so forth. The texts in the sample find evidence for this, as well as for transnational dimensions to the spread of violence. At the local level, Pierskalla and Hollenbach (2013) discover that the existence of conflict in a neighbouring spatial unit increases the likelihood of violence erupting in a given spatial unit. Such spatial dependency is not surprising, due to proximity between adjacent cities or regions. Additionally, as spatial units are overlaid over maps somewhat arbitrarily, such as in form of a grid-pattern, they could easily divide villages or regions sharing a common grievance or context, resulting in the observed effect.

More interestingly, Weidmann (2015) explores whether informational links created by ICTs can account for the spread of ethnic conflict across borders. He considers his paper as one of the first empirical tests of information linkages and their role in promoting conflict diffusion, largely independent of geography, but adding an additional layer of global interdependence. According to Weidmann (2015), studying

patterns of radio density (or mobile phone coverage for that matter) cannot say anything about information consumption, whereas fine-grained data on peer-to-peer communication is more conclusive although it has received little empirical attention. This lack of granularity is indeed a central limitation to most other studies in scope. Weidmann (2015) finds a strong positive effect of international phone calls on violent collective action, attesting to the importance of transnational links, including diaspora politics, as a strong factor affecting the risk of violence. There is evidence for a reinforcing effect of global informational linkages and geography, as "[a] country with conflict both in its geographic neighbour and in one of its communication partners has almost a seven-fold risk of conflict [erupting]" (Weidmann, 2015, p. 291).

This finding correlates with some of the research one of the authors has been involved within Ethiopia. There are, for example, a large number of Ethiopians living in diaspora across other African states, as well as internationally (United Nations Population Division, 2020). In past periods of protest and violent outbursts, the Ethiopian diaspora has been highly influential, galvanizing international attention and impacting domestic politics from afar (BBC News, 2016; Hairsine, 2020). This influence also extends to the dissemination of hateful speech and the sowing of ethnic divisions, where diasporic communities have been found to be actively involved, including during previous elections where we found a significant portion of online hate being propagated by diaspora communities in Norway (Gagliardone et al., 2016).

Conclusion

Through exploring the evidence around the role of ICTs, including social media platforms, in violent conflict in Africa, we have also sought to interrogate some of the prevailing myths and assumptions driving the headlines about new technologies and conflict, or peace. This approach follows others who have similarly challenged assumptions about how Africa might be portrayed whether in western media or in academic literature. We have sought to critically reflect on the narrative that violence in Africa is waiting to happen and (almost differently from other places in the world) online hate speech will be the spark. Others have noted how, in western media the continent is often overly generalized, or as lumped together as one rather than recognized for its vast diversity. And conflicts are framed as "ancient tribal hatreds" or tragedies "signifying darkness and hopelessness" rather than engaging with structural or historical challenges (Chari, 2010, p. 333).[6] At the same time, the actual dominance or pervasiveness of these narratives has become a recent focus for those highlighting the "myth" of representations of Africa (Scott, 2017) suggesting that

[6] There has, for example, been a focus on how the UK or US media has focused on repeating these stereotypes (Wall, 2009; Moeller, 1999) or, more recently, how China has sought to offer a counter-narrative to the destitute and violent continent, shoring up comradery with African states that seek an alternative to western-defined discourses (Gagliardone, 2013).

there is insufficient evidence to support these widespread generalizations (Nothias, 2018). We have similarly interrogated whether there is a myth, often in the western media, about how social media may be driving violence in Africa.

The trends across the literature examined offer some, albeit limited, insights into how online speech may relate to offline violence. The challenge has been that mechanisms which translate between technology and violence cannot be pinpointed. Articles make statements, describe a connection, but caveat that their results "do not reveal the exact causal mechanism in operation or the role of individual-level behaviour" (Pierskalla & Hollenbach, 2013, p. 221). In fact, most articles acknowledge that "the mechanisms of impact are also poorly understood" (Manacorda & Tesei, 2020, p. 535). Consequently, the bulk of available research describes evidence which may be considered circumstantial. While most studies incorporate control variables in their models, such as for the state of the economy (Manacorda & Tesei, 2020; Ackerman et al., 2021), thereby isolating individual moving parts, their results nevertheless do not probe the actual mechanisms, and do not reach the granular level.

One challenge with evidence reviews, as pointed out earlier, is that they tend to pick up quantitative studies as 'evidence' rather than qualitative studies. The disparity raised by the number of texts that made it through to our final sample show the need for greater breadth of qualitative research on this area. Statistical methodologies assume that causal mechanisms are independent and additive, which is unlikely (Dafoe & Lyall, 2015). They are consequently limited in what they can capture and are less sensitive to nuance than qualitative analyses. However, it is crucial to capture the interplay and messiness of factors and mechanisms with richer methodologies. Multiple authors emphasize the "complexity and context-dependence of technological effects" (Warren, 2015, p. 307) and call for richer data embedded in the communities to be studied. The trade-off may be that generalizable statements may not be possible, but the same could be said for the meta-level studies examined here.

As part of our ConflictNet project we have been working to fill some of these gaps. For example, we have sought to address the question of depth by focusing on government decision-making around internet shutdowns - often justified on national security grounds as a means of shutting down platforms that may have insufficient content moderation (particularly in African languages and in peripheral markets for Big Tech) and be seen as central vectors in spreading hate speech (such as in Ethiopia, Chad, or Cameroon) (Stremlau & Dobson, 2022; Marchant & Stremlau, 2020). At the same time, we have worked to complement this top-down data from elites in policymakers with more grounded ethnographic field work in communities such as Shashame, Ethiopia, that have been deeply affected by ongoing violence that has been associated with the spread of hate speech on social media platforms, and subsequently affected with internet shutdowns. Combining this range of empirical evidence has enabled us to qualitatively examine the locally perceived impacts of online hate but it is, as noted in our earlier limitations, very much context-specific and bounded by time, geography, and local politics. We have also sought breadth in our

research by experimenting with new ways of disseminating surveys, particularly in regions that may be conflict affected. By launching a Qualtrics survey using targeted advertisements on Facebook we have successfully polled thousands of social media users across eight sub-Saharan African countries about concepts of hate speech, whether and how they report hateful information online, and what impacts they believe such speech has on real world events. Again, we are well aware of the limitations of this approach but we believe its contribution – particularly in terms of its reach and comparative value - will significantly address some of the gaps we have highlighted in the evidence.

As more research is published on current conflicts such as Ethiopia, Sudan, or further afield (and beyond the scope of this article) such as Myanmar, it is expected that it will add nuance to many of the assumptions about media effects, particularly in Africa, and can push back on assumptions about the hypodermic needle effect - the assumptions that hate speech 'culminates' in events such as the war in Ethiopia. Such research will not only add to our understandings about social media but also power, economic, and social factors in violent conflict.

References

Absher, S., & Grier, K. (2019). *Can you hear me now? Good?? The Effect of Mobile Phones on Collective Violent Action in the Libyan Revolution* (No. 92627). Munich Personal RePEc Archive. https://mpra.ub.uni-muenchen.de/92627/

Ackermann, K., Awaworyi Churchill, S., & Smyth, R. (2021). Mobile phone coverage and violent conflict. *Journal of Economic Behavior & Organization*, *188*, 269–287. https://doi.org/10.1016/j.jebo.2021.04.038

Akçam, T. (2012). *The Young Turks' Crime against Humanity: The Armenian Genocide and Ethnic Cleansing in the Ottoman Empire*. Princeton University Press.

Akinwotu, E. (2021, 7 October). Facebook's role in Myanmar and Ethiopia under new scrutiny. *The Guardian*. https://www.theguardian.com/technology/2021/oct/07/facebooks-role-in-myanmar-and-ethiopia-under-new-scrutiny

Aouragh, M., & Alexander, A. (2011). The Arab Spring| The Egyptian Experience: Sense and Nonsense of the Internet Revolution. *International Journal of Communication*, *5*, 1344-1358. https://ijoc.org/index.php/ijoc/article/view/1191

Bailard, C. S. (2015). Ethnic conflict goes mobile: Mobile technology's effect on the opportunities and motivations for violent collective action. *Journal of Peace Research*, *52*(3), 323–337. https://doi.org/10.1177/0022343314556334

BBC News. (2016, November 24). Why is the Ethiopian diaspora so influential? https://www.bbc.com/news/world-africa-38076389

BSR. (2018). *Human Rights Impact Assessment: Facebook in Myanmar*. https://about.fb.com/wp-content/uploads/2018/11/bsr-facebook-myanmar-hria_final.pdf

Castells, M. (2011). *The rise of the network society*. John Wiley & Sons.

Chari, T. (2010). Representation or misrepresentation? The *New York Times's* framing of the 1994 Rwanda genocide. *African Identities, 8*(4), 333-349. https://doi.org/10.1080/14725843.2010.513242

Dafoe, A., & Lyall, J. (2015). From cell phones to conflict? Reflections on the emerging ICT–political conflict research agenda. *Journal of Peace Research*, *52*(3), 401–413. https://doi.org/10.1177/0022343314563653

Danning, G. (2018). Did Radio RTLM Really Contribute Meaningfully to the Rwandan Genocide?: Using Qualitative Information to Improve Causal Inference from Measures of Media Availability. *Civil Wars*, *20*(4), 529–554. https://doi.org/10.1080/13698249.2018.1525677

De Gregorio, G. & Stremlau, N. (2021). Information interventions and social media. *Internet Policy Review*, *10*(2), 1-25. https://doi.org/10.14763/2021.2.1567

De Gregorio, G. & Stremlau, N. (2023). Inequalities and Content Moderation. *Global Policy*, *14*, 870–879. https://doi.org/10.1111/1758-5899.13243

Department for International Development [DFID]. (2014). *Assessing the Strength of Evidence*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf

Des Forges, A. (2002). Silencing the Voices of Hate in Rwanda. In M. E. Price & M. Thompson (Eds.), *Forging Peace: Intervention, Human Rights and the Management of Media Space* (pp. 236–258). Edinburgh University Press.

Economist (2021, December 9). Can Facebook be blamed for pogroms against Rohingyas in Myanmar? Lawsuits in America and Britain seek billions of dollars in damages. *Economist*. https://www.economist.com/asia/2021/12/09/can-facebook-be-blamed-for-pogroms-against-rohingyas-in-myanmar

Fokou, G., Yamo, A., Kone, S., Koffi, A. J. D. A., & Davids, Y. D. (2022). Xenophobic violence in South Africa, online disinformation and offline consequences. *African Identities*, 1-20. https://doi.org/10.1080/14725843.2022.2157245

Fotheringham, A. S., & Wong, D. W. S. (1991). The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. *Environment and Planning A: Economy and Space*, *23*(7), 1025–1044. https://doi.org/10.1068/a231025

Fujii, L. A. (2011). *Killing Neighbours: Webs of Violence in Rwanda*. Cornell University Press. https://doi.org/10.7591/9780801458613

Gagliardone, I. (2013). China as a Persuader: CCTV Africa's First Steps in the African Mediasphere. *Ecquid Novi: African Journalism Studies*, *34*(3), 25–40. https://doi.org/10.1080/02560054.2013.834835

Gagliardone, I., Pohjonen, M., Beyene, Z., Zerai, A., Aynekulu, G., Bekalu, M., Bright, J., Moges, M. A., Seifu, M., Stremlau, N., Taflan, P., Gebrewolde, T. M., & Teferra, Z. (2016). *Mechachal: Online debates and elections in Ethiopia - from hate speech to engagement in social media*. Oxford University Press. https://doi.org/10.2139/ssrn.2831369

Global Witness. (2022, March 20). *Facebook approves adverts containing hate speech inciting violence and genocide against the Rohingya*. https://www.globalwitness.org/en/campaigns/digital-threats/rohingya-facebook-hate-speech/

Gohdes, A. R. (2018). Studying the Internet and Violent Conflict. *Conflict Management and Peace Science*, *35*(1), 89–106. https://doi.org/10.1177/0738894217733878

Grzyb, A. (2019, April 1). Debate continues about the media's role in driving Rwanda's genocide. *The Conversation*. http://theconversation.com/debate-continues-about-the-medias-role-in-driving-rwandas-genocide-114512

Hairsine, K. (2020, July 17). Ethiopia's influential diaspora split on Abiy. *Deutsche Welle*. https://www.dw.com/en/ethiopias-influential-diaspora-split-on-abiy/a-54209290

Herf, J. (2006). *The Jewish Enemy: Nazi Propaganda during World War II and the Holocaust*. Harvard University Press.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., & Welch, V.A. (Eds.) (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons.

Hollenbach, F. M., & Pierskalla, J. H. (2017). A re-assessment of reporting bias in event-based violence data with respect to cell phone coverage. *Research & Politics*, *4*(3). https://doi.org/10.1177/2053168017730687

Human Rights Council. (2018). *Report of the independent international fact-finding mission on Myanmar* (Report No. A/HRC/39/64). United Nations Human Rights Office of the High Commissioner. https://www.ohchr.org/sites/default/files/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_64.pdf

Ibreck, R., & de Waal, A. (2022). Introduction: Situating Ethiopia in Genocide Debates. *Journal of Genocide Research*, *24*(1), 83–96. https://doi.org/10.1080/14623528.2021.1992920

Ingrams, A. (2015). Mobile phones, smartphones, and the transformation of civic behavior through mobile information and connectivity. *Government Information Quarterly*, *32*(4), 506–515. https://doi.org/10.1016/j.giq.2015.07.007

International Telecommunication Union. (2019). *Individuals using the Internet (% of population)—Ethiopia*. The World Bank. https://data.worldbank.org/indicator/IT.NET.USER.ZS?end=2019&locations=ET&start=1990&view=chart

Kemp, S. (2021). *Digital in Ethiopia: All the Statistics You Need in 2021*. DataReportal. https://datareportal.com/reports/digital-2021-ethiopia

Kissi, E. (2021, August 2). Africa can prevent Ethiopia from going down Rwanda's path: here's how. *The Conversation*. https://theconversation.com/africa-can-prevent-ethiopia-from-going-down-rwandas-path-heres-how-165254

Li, D. (2004). Echoes of Violence: Considerations on Radio and Genocide in Rwanda. *Journal of Genocide Research*, *6*(1), 9–27. https://doi.org/10.1080/1462352042000194683

Loader, B.D., Vromen, A., & Xenos, M.A. (2014). The networked young citizen: social media, political participation and civic engagement. *Information, Communication & Society*, *17*(2), 143-150. https://doi.org/10.1080/1369118X.2013.871571

Macías-Medellín, M., & Atuesta, L. H. (2021). Constraints and military coordination: How ICTs shape the intensity of rebel violence. *International Interactions*, *47*(4), 692–719. https://doi.org/10.1080/03050629.2021.1898959

Manacorda, M., & Tesei, A. (2020). Liberation Technology: Mobile Phones and Political Mobilization in Africa. *Econometrica*, *88*(2), 533–567. https://doi.org/10.3982/ECTA14392

Marchant, E., & N. Stremlau (2020). A Spectrum of Shutdowns: Reframing Internet Shutdowns From Africa. *International Journal of Communication*, *14*, 4327-4342. https://ijoc.org/index.php/ijoc/article/view/15070/3188

McDoom, O. S. (2012). The Psychology of Threat in Intergroup Conflict: Emotions, Rationality, and Opportunity in the Rwandan Genocide. *International Security*, *37*(2), 119–155. https://doi.org/10.1162/ISEC_a_00100

McDoom, O. S. (2013). Who killed in Rwanda's genocide? Micro-space, social influence and individual participation in intergroup violence. *Journal of Peace Research*, *50*(4), 453–467. https://doi.org/10.1177/0022343313478958

Metzl, J. F. (1997). Rwandan Genocide and the International Law of Radio Jamming. *The American Journal of International Law*, *91*(4), 628–651. https://doi.org/10.2307/2998097

Mironko, C. (2007). The Effect of RTLM's Rhetoric of Ethnic Hatred in Rural Rwanda. In A. Thompson (Ed.), *The Media and the Rwanda Genocide* (pp. 125-135), Pluto Press.

Moeller, S. (1999). *Compassion Fatigue: How the Media Sell Disease, Famine, War, and Death*. Routledge.

Müller, K., & Schwarz, C. (2020). Fanning the Flames of Hate: Social Media and Hate Crime. *Journal of the European Economic Association, 19*(4), 2131–2167. https://doi.org/10.1093/jeea/jvaa045

Mutahi, P., & Kimari, B. (2017). *The Impact of Social Media and Digital Technology on Electoral Violence in Kenya* (IDS Working Paper No. 493). https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/13159/Wp493_Online.pdf?sequence=3

Njuguna, C. W., Gikandi, J., Kathuri-Ogola, L., & Kabaria-Muriithi, J. (2020). Social Media as a Battleground for Electoral Violence among Netizens in Africa: The Case of Kenyan Youth. *East African Journal of Arts and Social Sciences, 2*(1), 104-115. https://doi.org/10.37284/eajass.2.1.209

Nothias, T. (2018). How Western Journalists Actually Write About Africa: Re-assessing the Myth of Representations of Africa. *Journalism Studies*, *19*(8), 1138-1159. https://doi.org/10.1080/1461670X.2016.1262748

Olson, E. (2022, December 17). *A $1.6 billion lawsuit alleges Facebook's inaction fuelled violence in Ethiopia*. National Public Radio [NPR]. https://www.npr.org/2022/12/17/1142873282/facebook-meta-lawsuit-ethiopia-kenya-abrham-amare

Openshaw, S. (1984). Ecological Fallacies and the Analysis of Areal Census Data. *Environment and Planning A: Economy and Space*, *16*(1), 17–31. https://doi.org/10.1068/a160017

Pelican, M., Schumann, K., Plücken, S., & Drew, D. (2022). *Mbororo under Attack: Extreme Speech and Violence in the Anglophone Conflict in Cameroon* (Arnold-Bergstraesser-Institute Working Paper No. 21). https://www.arnold-bergstraesser.de/sites/default/files/field/pub-download/_abi_workingpaper_21_1.pdf

Pettersson, T., Davies, S., Deniz, A., Engström, G., Hawach, N., Högbladh, S., Sollenberg, M., & Öberg, M. (2021). Organized violence 1989–2020, with a special emphasis on Syria. *Journal of Peace Research*, *58*(4), 809–825. https://doi.org/10.1177/00223433211026126

Phys.org. (2015). *Facebook Founder in Colombia to Promote Affordable Internet*. https://phys.org/news/2015-01-facebook-founder-colombia-internet.html

Pierskalla, J. H., & Hollenbach, F. M. (2013). Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa. *The American Political Science Review*, *107*(2), 207–224. https://doi.org/10.1017/S0003055413000075

Posetti, J., Shabbir, N., Maynard, D., Bontcheva, K., & Aboulez, N. (2021). *The Chilling: Global trends in online violence against women journalists*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000377223

Power, S. (2001, September 1). Bystanders to Genocide. *The Atlantic*. https://www.theatlantic.com/magazine/archive/2001/09/bystanders-to-genocide/304571/

Rubin, M. (2021, October 3). Will Ethiopia's genocide be worse than Rwanda's? *Washington Examiner*. https://www.washingtonexaminer.com/opinion/will-ethiopias-genocide-be-worse-than-rwandas

Schoemaker, E., & Stremlau, N. (2014). Media and conflict: An assessment of the evidence. *Progress in Development Studies*, *14*(2), 181-195. https://doi.org/10.1177/1464993413517790

Scott, M. (2017). The Myth of Representations of Africa. *Journalism Studies*, *18*(2), 191-210. https://doi.org/10.1080/1461670X.2015.1044557

Stork, C., Calandro, E., & Gillwald, A. (2013). Internet going mobile: Internet access and use in 11 African countries. *Info*, *15*(5), 34–51. https://doi.org/10.1108/info-05-2013-0026

Straus, S. (2007). What Is the Relationship between Hate Radio and Violence? Rethinking Rwanda's "Radio Machete". *Politics & Society*, *35*(4), 609–637. https://doi.org/10.1177/0032329207308181

Stremlau, N. & Dobson, N. (2022). Information Controls and Internet Shutdowns in African Elections. *Journal of African Elections, 21*(2), 1-22. https://doi.org/10.20940/jae/2022/v21i2a1

Stremlau, N. (2018). *Media, Conflict and the State in Africa*. Cambridge University Press. https://doi.org/10.1017/9781108551199

Su, S. (2018, August 15). *Update on Myanmar*. Meta. https://about.fb.com/news/2018/08/update-on-myanmar/

Thompson, M. (1999). *Forging War: The Media in Serbia, Croatia, Bosnia and Hercegovina* (2 ed.). University of Luton Press.

Thompson, M., & Price, M. E. (2003). Intervention, Media and Human Rights. *Survival*, *45*(1), 183-202.

Towey, H. (2021, October 14). Facebook is fighting to keep records of its own investigation into the genocide of Rohingya Muslims in Myanmar out of court. *Business Insider*. https://www.businessinsider.com/facebook-mark-zuckerberg-myanmar-rohingya-genocide-case-whistleblower-2021-10

Trujillo, H. R., Elam, D., Shapiro, G., & Clayton, M. (2014). The Role of Information and Communication Technology in Preventing Election-Related Violence in Kenya, 2013. *Perspectives on Global Development and Technology*, *13*(1–2), 111–128. https://doi.org/10.1163/15691497-12341292

United Nations Population Division. (2020). *International Migrant Stock 2020*. https://www.un.org/development/desa/pd/content/international-migrant-stock

United Nations. (2019). *Hate speech 'dehumanizes individuals and communities: Guterres*. United Nations Outreach Programme on the Holocaust. https://www.un.org/en/outreach-

programme-holocaust/hate-speech-%E2%80%98dehumanizes-individuals-and-communities%E2%80%99-guterres

Wall, M. (2009). Africa on YouTube: Musicians, Tourists, Missionaries and Aid Workers. *The International Communication Gazette*, *71*(5), 393-407. http://doi.org/10.1177/1748048509104988

Warren, T. C. (2015). Explosive connections? Mass media, social media, and the geography of collective violence in African states. *Journal of Peace Research*, *52*(3), 297–311. https://doi.org/10.1177/0022343314558102

Weidmann, N. B. (2015). Communication networks and the transnational spread of ethnic conflict. *Journal of Peace Research*, *52*(3), 285–296. https://doi.org/10.1177/0022343314554670

Weidmann, N. B. (2016). A Closer Look at Reporting Bias in Conflict Event Data. *American Journal of Political Science*, *60*(1), 206–218. https://doi.org/10.1111/ajps.12196

Weidmann, N. B., Benitez-Baleato, S., Hunziker, P., Glatz, E., & Dimitropoulos, X. (2016). Digital discrimination: Political bias in Internet service provision across ethnic groups. *Science*, *353*(6304), 1151-1155. https://www.science.org/doi/abs/10.1126/science.aaf5062

Wilmot, C., Tveteraas, E., & Drew, A. (2021). *Dueling Information Campaigns: The War Over the Narrative in Tigray*. Media Manipulation Casebook. https://mediamanipulation.org/case-studies/dueling-information-campaigns-war-over-narrative-tigray

Yanagizawa-Drott, D. (2014). Propaganda and Conflict: Evidence from the Rwandan Genocide. *The Quarterly Journal of Economics*, *129*(4), 1947–1994. https://doi.org/10.1093/qje/qju020