

18 State-sponsored disinformation, hate speech, and violence

Mapping conceptual connections through Iran's anti-Bahá'í propaganda¹

Andrés Shoai and Sergio García Magariño

Introduction

Campaigns that combine hate speech and disinformation against vulnerable minorities persist in the 21st century as a strategy intimately connected to direct violence and frequently sponsored by specific governments. The role played by these discursive practices in the former Yugoslavia and Rwanda genocides has been well-documented (Oberschall, 2000; Yanagizawa-Drott, 2014), whereas the recent human rights catastrophe of the Rohingya minority in Myanmar provides an example of their persisting force well into the present century (Kironka & Peng, 2021; Ronan, 2019). In these cases, as in less notorious ones, verbal attacks and distortions of facts have been used effectively to incite violence and discrimination against specific populations. Additionally, the systematic use of online media for manipulation campaigns has been detected in 70 countries, according to an Oxford study that also highlights how digital spaces are co-opted by many authoritarian regimes (Bradshaw & Howard, 2019).

Aggressive language and deception are certainly not new, and neither is the fundamental contribution they can make to structural and direct violence. However, the concepts of “hate speech” and “disinformation” have only gained prominence in the last decade in scholarly work, legal frameworks, and policy debates (Kapantai et al., 2021; Paz et al., 2020). These works address the complexity, speed, global reach, and the loosening of ethical standards that characterise communications in the current media environment. Although a diverse range of studies about these two concepts is rapidly accumulating, the relationship between both notions—as well as their connection to state-sponsored behaviour—still constitutes a fuzzy subject.

In this chapter, we argue that developing conceptual clarity about this subject, and studying it empirically, are two important, pending, and interconnected tasks. On the one hand, by implementing a review of key theoretical developments and research works, this chapter will identify and describe some fundamental connections between state-sponsored disinformation and

state-sponsored hate speech, as well as the relationship of these communication practices with more flagrant violations of human rights against specific populations. On the other hand, to test this general framework and draw new insights from factual experience, those conceptual categories will be used in the process of organising and interpreting available information about a contemporary case, which is the persecution carried out by the Islamic Republic of Iran against the Bahá'ís, the largest non-Muslim religious minority in that country.

Naturally, disinformation and hate speech do not exist or thrive in a vacuum. A number of studies present some key contextual reference points, such as Wardle and Derakhshan's (2017) "information disorder", Bennet and Livingston's (2018) "disinformation order", and Chadwick's (2019) "crisis of public communication". What these assessments have in common is the observation of deteriorating democratic values in a rapidly-changing media environment—"a complex web of motivations for creating, disseminating and consuming ... 'polluted' messages" (Wardle & Derakhshan, 2017, p. 4), and the erosion of "authenticity, rationality, tolerance, and trust" in the dynamics of public opinion formation (Chadwick, 2019, p. 4). It is within this general context that we approach the following sections.

The disinformation-hate-violence triangle

Before exploring the entanglement between disinformation, hate speech, and violence, presenting a separate definition for each component of this "triangle" will prove useful. Although they have been conceptualised in various ways, a certain gravitation among scholars towards some key defining features for each one of these subjects constitutes a positive exception in the all-too-fragmented sphere of social studies.

Concerning disinformation, a succinct and widely accepted definition is "all forms of false, inaccurate, or misleading information designed, presented, and promoted to intentionally cause public harm or for profit" (European Commission, 2018, p. 3).

We understand hate speech, on the other hand, as "any kind of communication in speech, writing or behaviour that attacks or uses pejorative or discriminatory language with reference to a person or a group" on the basis of some "identity factor", such as race, religion, gender, or other (UN, 2019a, p. 2). Even when morally unacceptable or socially harmful, acts that fit under this wide definition of hate speech do not necessarily constitute a crime, especially in legislations where freedom of expression is considered a fundamental value. If speech assumes the form of "incitement", however, things change. The International Covenant on Civil and Political Rights (1966), for example, states that "any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law" (Art. 20).

This leads us to the third component of our conceptual triangle, which is the subject of violence. Galtung's (1990) classical theorisation on interrelated forms of violence provides important contextual elements about this topic (Filibeli & Ertuna, 2021). Whereas direct violence is composed of visible events, its variations across multiple circumstances can be explained in terms of its root causes in cultural and structural forms of violence. Under this perspective, physical or verbal aggression (direct violence) can be interpreted, for example, as emanating from the normalisation of hate speech, prejudices, and stereotypes against certain groups (cultural violence), which can in turn be institutionally sanctioned through state-sponsored social and economic discrimination (structural violence). At the same time, direct expressions of violence reinforce structural and cultural aspects.

Although a universally accepted framework that determines when a state is engaging in acts of disinformation, hate, or violence does not exist, it is important to note that the collective construction that comes closest to such a framework is the human rights principles and norms. The intimate relationship between human rights and the theme of this chapter will be evidenced in different ways in the following sections. These sections will seek to conceptually integrate state-sponsored expressions of hate speech and disinformation with more direct forms of violence that constitute flagrant violations of human rights. As we do not have records of previous theoretical works that address these subjects in an integrated way, our objective required the study of literature reviews that present the state of the art for each of these matters independently (in other words, for each "vertex" of the "triangle"), on the one hand, and publications that observe each of the possible relationships (or "sides" of the "triangle"), on the other. All the conceptual postulates and empirical observations that help define and interrelate these subjects were analysed by mutual comparison with the goal of classifying them into broad categories. Six statements emerged as a result; they will be presented as the components of a preliminary theoretical model that can guide future research efforts on the matter.

As mentioned in the introduction, our second specific objective is to analyse those connections through an empirical case study in order to explore how they operate in practice and generate new theoretical insights from social reality. The persecution of the Bahá'ís in Iran combines the three elements of state-sponsored hate, disinformation, and human rights violations, and can therefore provide paradigmatic value (Brookshaw & Fazel, 2008; Zabihi-Moghaddam, 2016). Complementing the academic literature, the case study will be developed through primary sources: official reports and resolutions from intergovernmental organisations, documents from the *Archives of Bahá'í Persecution in Iran*—which include official state and media documents from this country—, reports from human rights organisations, and accounts published by the Bahá'í International Community (BIC), the organisation that represents the Bahá'ís at the United Nations and other international fora. The documents will be

content-analysed through qualitative techniques using our theoretical model as a set of pre-defined organising categories of information.

Apart from defining conceptual postulates and observing a case in order to understand its internal features and mechanisms, our work intends to generate new insights for further theoretical refinement. This implies a dialogue between theory development and a case study, an epistemological strategy used in approaches such as *process tracing* (Bennett & Checkel, 2014).

Theoretical model

This section offers an initial theoretical model of six interrelated postulates to facilitate the analysis of state-sponsored hate and disinformation campaigns, based on the available conceptual and empirical studies on the matter. By “initial model” we mean an approximation to the object of study in order to capture its complexity and breadth by representing some of its key features and mechanisms, rather than a detailed network of explanatory interactions.

Hate relies on disinformation

While disinformation can exist and spread without relying on hate, the opposite is hardly imaginable. When hate drives action, truth and falsehood become relativised weapons. Although hate speech and disinformation seem to have a multifaceted relationship, this is the aspect that is predominantly assumed and reported in the literature. Terms like “invariably” (George, 2020, p. 146), “inextricably” (Kojan et al., 2020, p. 81), and “essential” (Holvoet, 2022, p. 2) are used to characterise the reliance of hate speech—including incitement—on disinformation. Previous research has found that partisan attacks, negativity, and hate speech are most likely to occur in false information that deviates the furthest from reality (Hameleers et al., 2022). It should be noted, however, that not all content imbued with hatred reveals its own aggressive characteristics. Hate speech can also appear to be “articulately and reasonably expressed” (Sorial, 2015, p. 299).

Hate incitement is a predictor of direct violence, including mass atrocities

Much work has been carried out to determine the relationship between incitement and violent action, and solid arguments about the former being a “precursor, indicator, predictor, and catalyst” of the latter have been made (Richter et al., 2018, p. 40). We have chosen the word “predictor” for this model because of its value for early warning and prevention purposes, and in order to move beyond debates about causality between speech and violence, as both are clearly multidimensional and complex objects of study.

It should also be noted that strong connections have been observed not only with respect to the concept of incitement but also when using the more general

concept of hate speech, with its intimate nexus to deception. “Hate speech begets hate crimes, as can misinformation and disinformation” (UN, 2021a, p. 7). In the case of mass atrocities, including genocide, emphasis should be added. The UN framework to analyse these crimes includes “acts of incitement or hate propaganda” targeting particular groups or individuals among the “triggering factors” (UN, 2019b, p. 17). Genocides involve the participation of large numbers of ordinary individuals transformed by “messages, imagery, and power relationships that dehumanize the intended targets” (Kopel, 2016, p. 452).

The seemingly ubiquitous presence of aggressive and deceiving discourses in online environments might suggest—at least to the uninformed observer—that such practices are somehow “diluted” across cyberspace, equally affecting diverse segments of the population, with attackers and victims constantly exchanging roles. However, this is not the case. According to the UN Special Rapporteur on minority issues, three-quarters of hate speech cases around the world target specific minorities (UN, 2021a). Of course, targets are not always numerically defined. Concepts like “vulnerable groups” and “gendered disinformation and hate” help to address other defining characteristics of target populations (Juárez-Rodríguez, 2015; Judson, 2020).

Disinformation (and counter-disinformation) pose threats to human rights

While the previous component of our theoretical model highlights the danger that public expressions of hate can pose to people’s fundamental rights, disinformation itself—even when free from hate speech—shares the same characteristic. This is because it depends on practices that “infringe on the autonomy and dignity of the person” (Glassius & Michaelsen, 2018, p. 3795). An illustration of this point is how disinformation about the COVID-19 pandemic affected the right to public health (Ramírez-Bañuelos, 2021). As reported by a recent UN document, there is “growing evidence that disinformation tends to thrive where human rights are constrained, where the public information regime is not robust, and where media quality, diversity and independence is weak” (UN, 2021b, p. 2). In such contexts, counter-disinformation initiatives can be understood as a way of protecting human rights. However, it should also be noted that counter-disinformation can be used to justify human rights violations (Colomina et al., 2021). What becomes clear when considering this double-edged relationship is that human rights principles offer “a normative framework that should underpin responses” to disinformation (Jones, 2019, p. 2).

State agency implies a unique concentration of resources

Troll factories, controlled mass media, surveillance, restrictions on information access, special propaganda offices, as well as instrumentalised education systems, state-religion institutions, and co-opted public figures are some of the resources used by certain states to deceive and spread hate for specific

interests. Clearly, states are “the most prolific users of disinformation” for a reason (Gunatilleke, 2021)—they possess all the necessary capabilities. A relevant example is the massive use of coordinated fake social media accounts controlled by human, bot, and cyborg state actors to spread computational propaganda and disinformation, especially in critical moments of public life (Beskow & Carley, 2020; Bradshaw & Howard, 2019; Niblock et al., 2022; Zannettou et al., 2019).

Furthermore, the state’s potential for discriminatory and deceiving communications cannot be fully assessed without considering the role that policy, administration, and legislation—even constitutional texts—can have in allowing, expressing, and engendering hate and disinformation. For example, Rohingya Muslims in Myanmar are left stateless through the constitutionally accepted notion of “national races” (*taingyintha*) (Cheesman, 2017).

Hate campaigns are structured and underpinned by wider narratives

When analysing a hate campaign, attention tends to be drawn to speech acts that are extreme, but these expressions do not work in a vacuum. “Hate campaigns comprise multiple, layered, loosely interlocking messages, disseminated by different actors over years or decades” (George, 2020, p. 147). These messages have some basic features, such as the dehumanisation of a target group and the reinforcement of a positive in-group feeling (Ibrahim, 2019; Kojan et al., 2020; Uyheng et al., 2022). A feature we mentioned previously is their reliance on disinformation narratives, which in turn are not completely out of touch with reality—they “alter, doctor or manipulate” information (Hameleers, 2023, p. 8). At the same time, speech acts happen in an even wider discursive context. Master narratives, which cultivate a primary social identity, provide the backdrop and are regularly refreshed with contemporary examples from the news and other sources (Levinger, 2018).

The strategies are multilevel and multichannel

Considering the scale and power of state structures, intentions to deceive and spread hate will usually move into action through specific objectives and multiple means and levels of implementation. These levels are conceptually organised in various ways, usually by interconnecting terms like operations, manoeuvres, tactics, behaviours, practices, and toolkits (e.g., Bhatia & Arora, 2024; European Commission, 2022; Lukito, 2020; Vargas et al., 2020). No model or terminology has been proven to be universally applicable. The important premise is to consider multiple strategic levels and relations organised around objectives or intentions.

On the other hand, as the object of study is communicative in nature and massive in its reach, the use of media is usually a key criterion when selecting research problems of disinformation and hate speech. Digitally mediated communication is the dominant focus, especially social media, but there is no reason

to assume that malicious agents will choose specific media types in the clear-cut manner that communication scholars frequently do. Usually, a strong communication objective calls for a multichannel approach. Neither can we assume that offline practices like a sermon from a pulpit or a pamphlet on a doorstep are intrinsically less relevant.

Case study: Iran’s hate and disinformation against the Bahá’ís

Adherents of the Bahá’í faith have faced violent opposition in Iran since their religion’s inception in the mid-nineteenth century, and they have been under a new wave of systematic and state-sponsored persecution from the early 1980s to the present day (Amanat, 2008; Ghanea, 2002; Milani, 2016, Zabihi-Moghaddam, 2016). With a community of around 350,000 members, Bahá’ís form the largest non-Muslim religious minority in that country. However, they are regarded by the state as “unprotected infidels” (UN, 2019c, p. 13). They are murdered with impunity, imprisoned without due process, their properties are confiscated, and their rights to work and education denied, just to mention some examples of violations.

This section will focus on the discursive aspect of the persecution by analysing state-sponsored disinformation and hate speech, as well as their relationship to violence. Considering the evident complexity of such a task, it is expected that the six components of our theoretical model will help us build a case study that provides both clarity and breadth, on the one hand, and generates new insights for theoretical reflection, on the other.

Reliance of hate on disinformation

Manipulation of truth has historically been a key ingredient in the “otherisation” narratives about the Bahá’ís in Iran (Chehabi, 2008; Yazdani, 2012), but the relation between the contemporary concepts of “hate speech” and “disinformation” requires scrutiny of specific communication contents. The organisation representing the Bahá’ís has gathered a sampling of official and semi-official anti-Bahá’í propaganda issued in Iran during a 16-month period, documenting around 400 articles, broadcasts, and other materials (BIC, 2011). The report identifies several recurring themes—that Bahá’ís are agents of Zionism or spies for Israel, that the Bahá’í Faith was created by imperialist powers like Great Britain or Russia, that it is a “misguided sect” associated with “cultlike” practices, and several others. It also observes a shift in these themes, which are “expanding from traditional theological attacks to those with a more contemporary flavor, with the goal of prejudicing the increasingly secular-minded Iranian population” (BIC, 2011, p. 13). The media examples provided by the report show that incitement can rely on disinformation to an extent where the distinction between the two becomes only analytical. For example, an article stating that Bahá’ís are, according to their teachings, “free to marry their daughters, sisters, aunts and

uncles” constitutes both an incitement to hatred and a baseless fabrication at the same time (BIC, 2011, p. 21).

The predicting qualities of hate speech

The Bahá’í case shows how aggressive speech precedes violence on many levels. For example, at a general or historical level, Ayatollah Khomeini’s discourse presented this minority as Iran’s “internal Other” for decades before climaxing in the Islamic Revolution (Yazdani, 2012), which in turn was followed by several atrocities—including the execution of over 200 Bahá’ís—by a regime that actively moved along the pathway towards genocide (Affolter, 2005; Bigelow, 1992; Momen, 2005).

An example of a more specific analytical focus is the Bahá’í International Community’s (2012) report about the happenings in Semnan between 2005 and 2012, where anti-Bahá’í seminars, sermons, pamphlets, and radio programmes created an atmosphere of animosity where both officials and citizens became free to act with impunity. The same approach can be narrowed down to an event-level of analysis—for example, as the document recounts, a single conference given in early December 2009 by the author of an anti-Bahá’í book was immediately followed by a series of raids in 20 Bahá’í homes (BIC, 2012, p. 17).

Counter-disinformation as a justification of human rights violations

In Iran, the precept of countering disinformation is instrumentalised by the government for Internet censorship and shutdowns, which are common in the country, and draft legislation for “preventing and countering publication of false information” has been created to increase control over the media (UN, 2021c, p. 12). However, direct attacks on communities and individuals, including Bahá’ís, are also made on the same grounds. A charge commonly used by Iranian officials is the “spreading of lies” (Sanasarian, 2012, p. 312). For example, court documents against Sahba Rezvani, who was imprisoned in 2008, charge her with “propaganda against the holy regime of the Islamic Republic of Iran” (BIC, 2012, p. 21).

The state’s concentration of resources

In Iran, “discrimination against the Baha’i community is legally sanctioned” (UN, 2017, p. 16) while “widely exercised by various organs of the Iranian state” (Milani, 2016, p. 137). Official budgets have included allocations for “educational” programmes to “confront” the Bahá’í Faith, and state organs have been established for that purpose (BIC, 2019, p. 12). The following example can help in understanding the institutionalised nature of hate. In 2007, the Education Department in Shiraz circulated a form to be completed by all non-Muslim

students. The section for “religion” listed only four options: “Christian”, “Jew”, “Zoroastrian”, and “Perverse Bahaist sect” (BIC, 2013, p. 23).

Hate and disinformation activities are implemented not only through such discriminatory policies, but also by blocking the application of non-discrimination policies. For example, Bahá'ís have long been denied access to any means of communication with the public and cannot counter the accusations propagated about them and their religion, which is in contradiction with Article 5 of Iran's Press Law (BIC, 2019).

A key dimension of the campaigns against the Bahá'ís, is a network of hundreds of political, judicial, and religious leaders that openly speak and write against this minority (BIC, 2018). Iran's Supreme Leader Ali Khamenei, for example, issued a religious decree (*fatwa*) on his website stating that “you should avoid any association and dealings with this perverse and misguided sect” (BIC, 2019, p. 13). The volume of anti-Bahá'í media content is difficult to measure, but it includes thousands of articles, videos, and other materials (<https://iranbahaipersecution.bic.org>). Bahá'í representatives have denounced that the propaganda is “shocking in its volume and vehemence, its scope and sophistication” (BIC, 2012, p. 2). They also inform that other sources of slander, such as graffiti, pamphlets, and anonymous letters, contain “without fail” the same language found in media affiliated with the government (BIC, 2019, p. 12).

Wider narratives

Hate campaigns are structured and underpinned by wider social narratives. In this respect, the case shows how such structures can persist over time, even when some of their protagonists and specific details can be “conveniently” supplanted in different periods. For example, the successive accusations of Bahá'ís as agents of the Russian, Ottoman, and British empires in different stages of Iranian history over the past 150 years have now taken on the form of Bahá'ís as agents of Zionism and Israel in contemporary official discourse (Tavakoli-Targhi, 2008).

At the same time, the master narrative about the Bahá'ís has a more fundamental religious component. As adherents of a post-Islamic religion, Bahá'ís are referred to as followers of “the misguided and misleading sect” (*fırqa-ye dālla-yi muđilla*) (Zabihi-Moghaddam 2016, p. 125). For this reason, they are considered religiously unclean (*najis*). This long-standing belief has persisted as a backdrop for anti-Bahá'í rhetoric and discrimination in the 21st century. For example, a *fatwa* signed by six Grand Ayatollahs in 2010 states that they “are even more *Najis* than dogs” (BIC, 2017, p. 120), and dogs are considered ritually impure in Islam. Using Ervin Staub's analysis of the road to genocide, Affolter (2005) shows how the anti-Bahá'í narrative is a fundamental strategy for excluding people from one's own “moral universe” (Affolter, 2005, p. 89).

The strategic perspective

Currently available information on the Bahá'í case shows the value of gathering evidence at the highest level of government in order to understand strategies. The official policy of the Iranian government against this minority is summarised in a secret memorandum obtained in 1993 by a UN Representative (BIC, 2017). Signed by the Supreme Leader, Ali Khamenei, it states that “the government’s dealings with them must be in such a way that their progress and development are blocked” (BIC, 2017, p. 95). This document, which remains in effect 30 years later, outlines measures to restrict the educational, economic, and cultural life of Iranian Bahá'ís, including the creation of special propaganda offices. On the other hand, public speeches and *fatwas* of the Supreme Leader set the tone for anti-Bahá'í rhetoric and are usually followed by multiple amplifying messages in media organisations affiliated with the government (BIC, 2012).

While massive disinformation operations deployed by the government of Iran through fake accounts on Twitter and other social media have been detected and analysed (Bradshaw & Howard, 2019; Nembr & Gangware, 2019; Niblock et al., 2022), the study of official documents, public speeches, and legacy media content can place social media activity within wider strategies and detect specific state policies and campaigns. In this context, offline activities cannot be underestimated. Visual arts festivals where participants are incentivised to design anti-Bahá'í posters, or the presentation of numerous anti-Bahá'í books at Tehran book exhibitions and fairs, are examples of such operations (BIC, 2022).

Discussion and conclusions

Describing a state-sponsored campaign of hate and disinformation, and its connections to violence, can be a difficult challenge. The present chapter has sought to facilitate this task by pointing out some of the key components and interconnections that make up such campaigns. Through a series of conceptual statements, we have shown how hate speech uses disinformation and constitutes a predictor of direct violence, including mass atrocities, and we have also highlighted how disinformation and alleged counter-disinformation actions—even when analysed independently from hate speech—can pose threats to human rights. Additionally, we have considered the characteristics that state sponsorship provides to this disinformation-hate-violence triangle—an incomparable concentration of resources and the execution of strategies through multiple channels and levels of implementation. Our model also points out the importance of understanding governmental communication strategies in contexts that go beyond state structures, as these discursive practices usually derive their power from wider social narratives that are in turn reinforced through state action.

The analytical value and the main precepts of this initial model were illustrated through Iran’s anti-Bahá'í propaganda. This case study has also generated some conceptual insights that, even when not fundamentally altering the

framework, can prove useful in future research endeavours. Concerning the reliance of hate speech on disinformation, the Bahá'í case shows how this relationship can be strong to an extent where the distinction between the two becomes only analytical, as deception and hate incitement can be constituent elements of one discourse, one campaign, and even one statement. Regarding the predictive qualities of hate speech in relation to violence, the case reveals that these qualities can be observed at different levels of analysis, ranging from historical processes to event-specific views.

With respect to the structural properties of hate speech, the connection of these discourses to wider social narratives, and the potency they can gain when sponsored by the state, the case study helps us move from general notions to more specific analytical categories that can be helpful in gathering and organising information. Furthermore, it shows how evidence from official documents, public speeches, state-affiliated legacy media, and offline communications can clarify the underlying strategies and provide context to the social media operations that are usually the focus of disinformation studies.

As this chapter focused on state-sponsored campaigns—and not on the wider social processes of hate speech and disinformation—some contextual aspects have been excluded from the analysis, such as the roles of non-state actors like social media platforms and users, or the response patterns shown by the populations under attack (Karlberg, 2010). Although our approach has intentionally chosen state action as the focal point, it is only based on an analytical distinction, as state action is connected to wider societal processes.

Note

- 1 Andrés Shoai's participation in this research was supported by a Juan de la Cierva Formación grant (FJC2021-0471157-I) funded by the Ministry for Science and Innovation of Spain (MCIN/AEI/10.13039/501100011033) and the European Union NextGenerationEU/PRTR. The work also received support from the "Media Repertoires and Practices in Adolescence and Youth: Uses, Cyber Well-being, and Digital Vulnerabilities on Social Networks" project (PID2022-138281NB-C21), Knowledge Generation Plan, State Research Agency of Spain.

Sergio García-Magariño's participation in this research is part of a Spanish National Project funded by the Ministry for Science and Innovation, "Smart War. Old wars and new technologies: a critical test for the regulation of political violence" (PDC2021-121472-I00), and the European Union Next Generation funds.

References

- Affolter, F. (2005). The specter of ideological genocide: The Baha'is of Iran. *War Crimes Genocide & Crimes against Humanity*, 1, 75–114.
- Amanat, A. (2008). The historical roots of the persecution of Babis and Baha'is in Iran. In D. Brookshaw & S. Fazel (Eds.), *The Bahá'ís of Iran: Socio-historical studies* (pp. 170–183). Routledge.

- Bahá'í International Community (BIC). (2011). *Inciting hatred: Iran's media campaign to Demonize Bahá'ís*. <https://www.bic.org/publications/inciting-hatredirans-media-campaign-demonize-bahais>
- Bahá'í International Community (BIC). (2012). *Inciting hatred: The Bahá'ís of semnan. A case study in religious hatred*. <https://www.bic.org/publications/inciting-hatred>
- Bahá'í International Community (BIC). (2013). *Violence with Impunity: Acts of aggression against Iran's Bahá'í community*. <https://www.bic.org/publications/violence-impunity>
- Bahá'í International Community (BIC). (2017). *The Bahá'í question revisited: Persecution and resilience in iran*. https://www.bic.org/sites/default/files/pdf/iran/tbqu_english_170913_001.pdf
- Bahá'í International Community (BIC). (2018, January). *Islamic Republic of Iran: Non-implementation of accepted and partially-accepted UPR recommendations*. https://www.bic.org/sites/default/files/pdf/iran/iran_upr_mid-term_report_january_2018.pdf
- Bahá'í International Community (BIC). (2019, July). *The Bahá'ís of Iran – A persecuted community*. https://www.bic.org/sites/default/files/pdf/iran/the_bahais_of_iran_-_a_persecuted_community.pdf
- Bahá'í International Community (BIC). (2022, July). *The Bahá'ís of Iran – A persecuted community*. https://www.bic.org/sites/default/files/pdf/iran/iran_overview_of_persecution-0722_.pdf
- Bennet, L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139.
- Bennett, A., & Checkel, J. (2014). *Process tracing: From metaphor to analytical tool*. Cambridge University Press.
- Beskow, D., & Carley, K. (2020). Characterization and comparison of Russian and Chinese disinformation campaigns. In K. Shu et al. (Eds.), *Disinformation, misinformation, and fake news in social media* (pp. 63–81). Springer.
- Bhatia, K., & Arora, P. (2024). Discursive toolkits of anti-muslim disinformation on Twitter. *The International Journal of Press/Politics*, 29(1), 253–272.
- Bigelow, K. (1992). A campaign to deter genocide: The Bahá'í experience. In H. Fein (Ed.), *Genocide Watch* (pp. 189–196). Yale University Press.
- Bradshaw, S., & Howard, P. (2019). *The global disinformation disorder: 2019 Global inventory of organised social media manipulation* (Working Paper). Project on Computational Propaganda.
- Brookshaw, D., & Fazel, S. (Eds.). (2008). *The Baha'is of Iran: socio-historical studies (1st ed.)*. Routledge.
- Chadwick, A. (2019). *The new crisis of public communication: Challenges and opportunities for future research on digital media and politics*. Loughborough University.
- Cheesman, N. (2017). How in Myanmar 'national races' came to surpass citizenship and exclude Rohingya. *Journal of Contemporary Asia*, 47(3), 461–483.
- Chehabi, H. (2008). Anatomy of prejudice. Reflections on secular anti-Baha'ism in Iran. In D. Brookshaw & S. Fazel (Eds.), *The Bahá'ís of Iran: Socio-historical studies* (pp. 184–199). Routledge.
- Colomina, C., et al. (2021). *The impact of disinformation on democratic processes and human rights in the world*. European Parliament.

- European Commission. (2018). *A multi-dimensional approach to disinformation. Report of the independent High-level Group on fake news and online disinformation to the European Commission*. <https://op.europa.eu/en/publication-detail/-/publication/6cef4df8b-4cea-11e8-be1d-01aa75ed71a1>
- European Commission. (2022). *The strengthened code of practice on disinformation*. <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>
- Filibeli, T., & Ertuna, C. (2021). Sarcasm beyond hate speech: Facebook comments on Syrian refugees in Turkey. *International Journal of Communication*, 15, 2236–2259.
- Galtung, J. (1990). Cultural violence. *Journal of Peace Research*, 27(3), 291–305.
- George, C. (2020). The scourge of disinformation-assisted hate propaganda. In M. Zimdars & K. McLeod (Eds.), *Fake news: Understanding media and misinformation in the digital age* (pp. 145–152). The MIT Press.
- Ghanea, N. (2002). *Human rights, the UN and the Bahá'is in Iran*. George Ronald.
- Glasius, M., & Michaelsen, M. (2018). Illiberal and authoritarian practices in the digital sphere—prologue. *International Journal of Communication*, 12, 3795–3813.
- Gunatilleke, G. (2021). Reframing the debate: The state & disinformation in Sri Lanka. *South Asia @LSE Blog*. <https://blogs.lse.ac.uk/southasia/2021/12/06/reframing-the-debate-the-state-disinformation-in-sri-lanka/>
- Hameleers, M. (2023). Disinformation as a context-bound phenomenon: Toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination. *Communication Theory*, 33(1), 1–10.
- Hameleers, M., van der Meer, T., & Vliegthart, R. (2022). Civilized truths, hateful lies? Incivility and hate speech in false information – evidence from fact-checked statements in the US. *Information, Communication & Society*, 25(11), 1596–1613.
- Holvoet, M. (2022). International criminal liability for spreading disinformation in the context of mass atrocity. *Journal of International Criminal Justice*, 20(1), 223–250.
- Ibrahim, A. (2019). Theorizing the journalism model of disinformation and hate speech propagation in a Nigerian democratic context. *International Journal of E-Politics (IJEP)*, 10(2), 60–73.
- International Covenant on Civil and Political Rights. (1966). <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>
- Jones, K. (2019). *Online disinformation and political discourse. Applying a human rights framework*. Chatham House.
- Juárez-Rodríguez, J. (2015). *Estrategias y campañas de desinformación gubernamental y manipulación informativa en relación a los feminicidios y secuestros de mujeres y niñas en Ciudad Juárez entre 1993 y 2013*. Doctoral Thesis, Universidad Complutense de Madrid.
- Judson, E., et al. (2020). Engendering hate: The contours of state-aligned gendered disinformation online. *Demos*. <https://www.ndi.org/publications/engendering-hate-contours-state-aligned-gendered-disinformation-online>
- Kapantai, E., et al. (2021). A systematic literature review on disinformation: Toward a unified taxonomical framework. *New Media & Society*, 23(5), 1301–1326.
- Karlberg, M. (2010). Constructive resilience: The Baha'i response to oppression. *Peace & Change*, 35(2), 222–257.

- Kironska, K., & Peng, N. (2021). How state-run media shape perceptions: An analysis of the projection of the Rohingya in the global new light of Myanmar. *South East Asia Research*, 29(1), 16–31.
- Kojan, L., et al. (2020, October 26–27). Defend your enemy. A qualitative study on defending political opponents against hate speech online. In M. van Duijn et al. (Eds.), *Disinformation in open online media, Proceedings of the second multidisciplinary international symposium* (pp. 80–94). Springer.
- Kopel, H. (2016). The case for sanctioning state sponsors of genocide incitement. *Cornell International Law Journal*, 49(2), 415–468.
- Levinger, M. (2018). Master narratives of disinformation campaigns. *Journal of International Affairs*, 71(1.5), 125–134.
- Lukito, J. (2020). Coordinating a multi-platform disinformation campaign: Internet research agency activity on three U.S. Social media platforms, 2015 to 2017. *Political Communication*, 37(2), 238–255.
- Milani, S. (2016). Situation of the Baha'i minority in Iran and the existing legal framework. *Journal of International Affairs*, 69(2), 137–148.
- Momen, M. (2005). The Babi and Baha'i community of Iran: A case of 'suspended genocide'? *Journal of Genocide Research*, 7(2), 221–241.
- Nemr, C., & Gangware, W. (2019). Weapons of mass distraction: Foreign state-sponsored disinformation in the digital age. *Park Advisors*. <https://www.park-advisors.com/disinfoforeport>
- Niblock, I., Wallis, J., & Zhang, A. (2022). *Understanding global disinformation and information operations*. ASPI International Cyber Policy Centre. https://www.aspi.org.au/report/understanding_global_disinformation_information_operations
- Oberschall, A. (2000). The manipulation of ethnicity: From ethnic cooperation to violence and war in Yugoslavia. *Ethnic and Racial Studies*, 23(6), 982–1001.
- Paz, M. A., Montero-Díaz, J., & Moreno-Delgado, A. (2020). Hate speech: A systematized review. *Sage Open*, 10(4), DOI: 10.1177/2158244020973022.
- Ramírez-Bañuelos, J. (2021). Repensar la desinformación pública en casos de emergencia en las democracias americanas. *Anuario de Derechos Humanos*, 17(2), 449–471.
- Richter, E., Markus, D., & Tait, C. (2018). Incitement, genocide, genocidal terror, and the upstream role of indoctrination: Can epidemiologic models predict and prevent? *Public Health Reviews*, 39(1), 1–22.
- Ronan, L. (2019). Extreme speech in Myanmar: The role of state media in the Rohingya forced migration crisis. *International Journal of Communication*, 13, 3203–3224.
- Sanasarian, E. (2012). Nationalism and religion in contemporary Iran. In A. Longva & A. Roald (Eds.), *Religious minorities in the middle east* (pp. 309–324). Brill.
- Sorial, S. (2015). Hate speech and distorted communication: Rethinking the limits of incitement. *Law and Philosophy*, 34, 299–324.
- Tavakoli-Targhi, M. (2008). Anti-bahaism and Islamism in Iran. In D. Brookshaw & S. Fazel (Eds.), *The Bahá'is of Iran: Socio-historical studies* (pp. 200–231). Routledge.
- United Nations. (2017). *General assembly. Human rights council. Report of the special rapporteur on the situation of human rights in the Islamic Republic of Iran, A/HRC/34/65*. United Nations.
- United Nations. (2019a). *Secretary-general. United Nations Strategy and Plan of Action on Hate Speech*. https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_EN.pdf

- United Nations. (2019b). *Office on genocide prevention and responsibility to protect. Framework of analysis for atrocity crimes. A tool for prevention*. United Nations.
- United Nations. (2019c). *General assembly. Human rights council. Report of the special rapporteur on the situation of human rights in the Islamic Republic of Iran, A/74/188*. United Nations.
- United Nations. (2021a). *General assembly. Human rights council. Report of the special rapporteur on minority issues, Fernand de Varennes, A/HRC/46/57*. United Nations.
- United Nations. (2021b). *General assembly. Human rights council. Report of the special rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan, A/HRC/47/25*. United Nations.
- United Nations. (2021c). *General assembly. Human rights council. Situation of human rights in the Islamic Republic of Iran. Report of the special rapporteur on the situation of human rights in the Islamic Republic of Iran, Javaid Rehman, A/76/160*. United Nations.
- Uyheng, J., Cruickshank, I., & Carley, K. (2022). Mapping state-sponsored information operations with multi-view modularity clustering. *EPJ Data Science*, 11(25), 1–22.
- Vargas, L., Emami, P., & Traynor, P. (2020, November). *On the detection of disinformation campaign activity with network analysis*. Proceedings of the 2020 ACM SIGSAC Conference on Cloud Computing Security Workshop, pp. 133–146.
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe Report. <http://tverezo.info/wp-content/uploads/2017/11/PREMS-162317-GBR-2018-Report-desinformation-A4-BAT.pdf>
- Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the Rwandan Genocide. *The Quarterly Journal of Economics*, 129(4), 1947–1994.
- Yazdani, M. (2012). The Islamic revolution’s internal other: The case of Ayatollah Khomeini and the Baha’is of Iran. *Journal of Religious History*, 36(4), 593–604.
- Zabihi-Moghaddam, S. (2016). State-sponsored persecution of Baha’is in the Islamic Republic of Iran. *Contemporary Review of the Middle East*, 3(2), 124–146.
- Zannettou, S., et al. (2019). Disinformation warfare: Understanding state-sponsored trolls on Twitter and their influence on the web. *WWW ’19: Companion Proceedings of The 2019 World Wide Web Conference*(pp. 218–226). Association for Computing Machinery.