

# Denunciation and doxing: towards a conceptual model of digital vigilantism

Daniel Trottier

To cite this article: Daniel Trottier (2019): Denunciation and doxing: towards a conceptual model of digital vigilantism, *Global Crime*, DOI: [10.1080/17440572.2019.1591952](https://doi.org/10.1080/17440572.2019.1591952)

To link to this article: <https://doi.org/10.1080/17440572.2019.1591952>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 25 Mar 2019.



Submit your article to this journal [↗](#)



Article views: 1113



View related articles [↗](#)



View Crossmark data [↗](#)

## Denunciation and doxing: towards a conceptual model of digital vigilantism

Daniel Trottier

Department of Media and Communication, Erasmus Universiteit Rotterdam, Rotterdam, Netherlands

### ABSTRACT

Individuals rely on digital media to denounce and shame other individuals. This may serve to seek justice in response to perceived offences, while often reproducing categorical forms of discrimination. Both offence taking and its response are expressed online by gathering and distributing information about targeted individuals. By seeking their own form of social and/or criminal justice, participants may supersede institutions and formal procedures. Yet digital vigilantism includes shaming and other forms of cultural violence that are not as clearly regulated. They may feed from state or press-led initiatives to shame targets, or simply to gather information about them. Digital vigilantism remains a contested practice: Terms of appropriate use are unclear, and public discourse may vary based on the severity of the offence, the severity of response, and on participants' identities and affiliations. This paper advances a conceptually informed model of digital vigilantism, in recognition of its coordinated, moral and communicative components. Drawing upon literature on embodied vigilantism as well as concurrent forms of online coordination and harassment, it considers recent cases in a global context in order to direct subsequent analysis of how digital vigilantism is rendered meaningful.

### KEYWORDS

Vigilantism; digital media; shaming; online justice

Current social and political discourse is heavily shaped through media practices in both obvious and not so obvious ways. In particular there are specific affordances and cultural expectations that shape contemporary digital media use, which direct our way of relating and coordinating among citizens. Because of the ease with which users can search, solicit and disseminate personal information, these can be used to mobilise outrage for various cultural and political ends. In some cases these operate as cohesive organisations, as when the Russian group Lev Protiv shame and call attention to individuals who seem publicly intoxicated. In other cases, high-profile criminal events mobilise civil society to identify suspects, as was the case following the Boston Marathon bombing in 2013. Still others are one-off incidents that nevertheless are contextualised, embedded in and rendered meaningful through contemporary media practices. As an example, in 2013 a conference attendee publicly denounced via Twitter two other attendees for their sexualised comments, and uploaded a photo of the perpetrators. One of the targets lost his job. Yet the denouncer, an African-American woman, also lost her job in consequence, and became the target of racist and sexist vitriol. Journalists and

**CONTACT** Daniel Trottier  [trottier@eshcc.eur.nl](mailto:trottier@eshcc.eur.nl)

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

other public figures picked up on the incident, furthering the visibility of all three primary participants, in order to advance public discourse about gendered harassment as well as 'call-out culture' as a viable form of social action.<sup>1</sup>

These examples provide a brief account of the range of social actors engaging in digitally mediated forms of vigilantism. This refers to a set of practices to scrutinise, denounce and even leverage harm against those deemed to transgress legal and/or moral boundaries, with the intention of achieving some form of justice.<sup>2</sup> These practices can be understood as facilitated by media affordances that enable users to share and distribute content with each other, coupled with socio-cultural expectations that serve to reproduce and even renegotiate what is considered acceptable within a particular social context.<sup>3</sup> The notion of digital vigilantism includes responses to criminal events as well as offences that transgress moral or normative boundaries. While high-profile cases include responses to criminal acts, the same tools and practices can be used in contexts where the justice sought is more social than legal or where are not aligned with normative demands.<sup>4</sup> In discussing embodied forms of vigilantism, Johnston notes the complexity of distinguishing legal from moral pursuits, citing cases featuring a confluence of these frameworks.<sup>5</sup> Likewise, Moncada notes that while it is possible to consider moral vigilantism as a variant of embodied vigilantism, this risks overextending such a concept, notably when other terms such as hate crime are more apt.<sup>6</sup> This is indeed true, and an important task in this article is to consider the conceptual overlap of neighbouring terms. Yet it remains important to consider that vigilantism serves to uphold institutionalised norms, and the status of these norms and the degree to which they are institutionalised remains contested.<sup>7</sup> For instance in and beyond the Anglo-American context citizens have directed public attention to longstanding forms of sexual assault and harassment. This involves unambiguous violations of laws, but also behaviour and attitudes that are insufficiently targeted in existing legal frameworks, yet remain actionable.<sup>8</sup>

While it has always been possible denounce others, digital vigilantism implicates the active and passive participation of other individuals as well as several types of organisations. Contemporary media cultures allow virtually anybody to engineer and exploit the visibility of a targeted individual for an assortment of social, cultural, economic and political ends. This is not only through the capture of original footage of a target, or the repurposing of content already available online, but also through the fabrication of image or video content made to appear veracious.<sup>9</sup> Despite the range of motivations and desired outcomes fuelling these practices, they are united through their denunciatory nature, in which they link particular grievances to targeted individuals. These occur in cultural contexts that seek and celebrate forms of social justice – or some form of shared values and morality more generally – through the vigilance and participation of an assembly of social actors. To be clear, these cultural practices can also serve to reproduce misogyny, racism and other forms of categorical discrimination. Such practices risk being overshadowed in public discourse by adjacent phenomena like hate speech and cyberbullying. Yet digital vigilantism is a pervasive and mobilising force, linking up with these harms, but also broader organisational logics that solicit and shape mediated activity.

Recent press coverage on weaponised media places emphasis on cyber security incidents such as the SONY/North Korea hacking incident in 2014, Ashley Madison in 2015, and of American and French presidential candidates more recently.<sup>10</sup> In doing so,

they risk overlooking more routinized and accessible practices. While hacking and other cyber threats can indeed harm the social standing of those who are exposed, they involve skillsets and technologies often limited to an elite few. Yet even high-tech forms of hacking enable further repurposing of this data through more vulgarised forms of data searching and distribution. Social harm might not only emanate from those who first access and source compromising data from a secure location, but also from those who apply the now-public data to particular contexts, or for particular audiences.

Digital vigilantism may be framed as the dark side of online engagement through related practices such as vitriol or trolling. Yet a key concern in this area of study is that digital vigilantism is not simply problematic or deviant. Rather, it may be understood as a standardised mode of communication, and may elicit popular support under specific conditions.<sup>11</sup> It is a mode of organising for individuals as well as institutions, by gaining and leveraging attention against targeted others. It is also a wilful strategy that is partly informed by (but also exceeding) tabloid culture and reality TV.<sup>12</sup> Likewise, social media platforms are reluctant to fully prohibit practices that are entirely dependent on their infrastructure. While mindful of social responsibility and the impact on their own reputations, platforms like YouTube and Twitter benefit from the frenzy of user activity emerging from denunciatory campaigns, notably as even critical engagements will occur through these platforms. As was the case with social media's initial emergence in the mid 2000s, digital vigilantism involves a set of reasonably accessible technologies as well as rapidly popularised cultural practices that are spread to a range of social actors. Social media was rendered culturally meaningful and negotiated in terms of perceived appropriateness in various cultural contexts.<sup>13</sup> Indeed, concepts such as privacy, reputation and even sharing evoke debates that are opened up anew under current conditions.

This paper will advance a comprehensive understanding of the use of visibility and reputation of individuals as a means of communicating, organising, and governing by digital media users. The next section considers recent scholarship related justice seeking online. This is followed by a consideration of how such practices are manifest discursively by principal social actors. This paper then offers a tentative procedural model to understand digital vigilantism, including a conceptualisation of core concepts and adjacent terms, along with a consideration of key questions for subsequent research.

## **Mediated justice seeking and its scholarly challenges**

Individuals can exploit the visibility and reputation of other individuals as a means of communicating. Targeted individuals may either voluntarily post content about themselves or their peers – though not necessarily consenting to the appropriation of that content – or may be involuntarily recorded engaged in untoward behaviour. This content becomes a means to express outrage about particular events, but also serves as convenient and tangible focal points for more abstract societal concerns. Moreover, it becomes a means to draw attention to both press outlets and social media platforms. Both collective outrage and the target's reputation are leveraged for a variety of ends. While scholarly approaches to online harassment contend that the invisibility of social actors enables abusive behaviour, digital vigilantism is predicated on the amplified visibility of the person receiving abuse.<sup>14</sup> The relative ease of access to a substantial audience is also a feature that is now available to those who initiate the shaming of

others. Moreover, these audiences make take on an active role in these campaigns by re-circulating damning content, adding their own commentary or even information about targets. In some cases, the audience can be minimal, but the imagined or anticipated social impact can still harm the targeted individual as well as the broader community to which they belong.<sup>15</sup> Yet in many cases, individuals are the subsequent targets of actualised harassment, as well as consequences that may impact their life chances, for example, if evidence of their offence emerges during an online search when seeking employment, or while held at a border crossing.<sup>16</sup>

Instances of digital vigilantism are typically characterised by spontaneous and unreflexive activity as well as coordinated actions. Embodied vigilantism necessitates a degree of planning, or is otherwise categorised a self-defence.<sup>17</sup> Yet affordances of social platforms such as their connectivity (van Dijck and Poell 2013) mean that planned coordination can follow spontaneous actions, and vice versa.<sup>18</sup> In other words some related practices may be spontaneous, such as immediately capturing and uploading content as well as commenting on or sharing that content. Yet other aspects require deliberate planning, such as managing a presence on a digital platform, or coordinating a denunciatory response. For states, but also media outlets, social media platforms and other organisations, such incidents may amount to a form of 'soft' governance, where social norms are enforced and renegotiated and expressed through digital media and through communities.<sup>19</sup> Given the ease through which various social actors can remediate content by re-posting it on any available social platform, it remains possible that an incident that one social actor first launches is then repurposed and appropriated by another, for example, when tabloid newspapers provide extensive coverage of uncivil behaviour first documented by an individual. Approaching this in terms of a single category of social actor is tempting, but overlooks a broader context, notably as they take up similar devices and practices and may even sustain symbiotic relations. When considering the recent history of using digital media to monitor and discipline one another, we may speak of a mutual augmentation of shaming, as participants demand audiences, trigger responses, and produce content that are circulated by others.<sup>20</sup>

Public discourses and social responses to denunciations and discrediting content vary tremendously, in part because assessments of these practices often have to reconcile events that vary radically in terms of ideology and intent. To some degree shaming and moralising can be socially progressive by raising awareness of social issues such as gendered forms of harassment, but are also used to reproduce privilege and asymmetrical power relations. Most cases emerge in response to an offensive act, and are often expressed in criminal, ethical and moral terms. Even high-profile instances of sexist and racist abuse such as Gamergate attempt to frame their actions in terms of a moral high ground by invoking a concern over ethics in video game journalism.<sup>21</sup> Likewise, some incidents occur in context of broader cultural shifts such as #metoo, while others fail to evoke an impact to the same degree.<sup>22</sup> What unites these incidents is the use of digital media to render offences, offenders – and in some cases, the denunciator themselves – visible to public scrutiny. These practices have some precedent, and to some degree this is captured through the use of the term vigilantism. This term stresses citizen-led interventions against other individuals, often overlapping with and even contesting institutional forms of justice seeking, while often reproducing established cultural values.

Digital vigilantism is made meaningful through a range of social actors who are either directly participating in or offering commentary on events. As such, the vocabulary employed has a formative role in terms of asserting and contesting what is considered acceptable. Many terms employed to describe practices are sourced from digital media subcultures ('ownage'; 'rekt'; 'btfo') and betray a juvenile and clandestine nature. Terms may highlight contemporary media affordances ('leaks'; 'screencapping') that underline the ease with which information can be transmitted. Terms such as 'receipts' (denoting a proof of offence such as a screenshot of a since-deleted message) implicate an economy and legitimisation of outrage. Some terms used may be platform specific, such as 'subtweeting' (posting a critical statement about someone without directly identifying them). Yet as these practices can easily spread beyond any single platform, the relevance of such terms may exceed the immediate context of (in the above example) Twitter. Participants may also be framed through subject positions such as social justice warriors, keyboard warriors or trolls, but also '*reaguurders*' in the Dutch context, and '*diaosi*' in China, each of which carry connotations that appear both self-deprecating and mobilising.<sup>23</sup> Editorials and blog posts with titles like 'We with the Pitchforks' (Mae 2016) suggest an embracing attitude towards mediated vigilantism (in this case, following the denunciation of a convicted rapist). Other concepts such as the 'vampire's castle' and 'weaponised autism' both refer to collective and contextualised attempts to inflict harm upon others. We may presume that denunciations serve to explicitly and implicitly communicate desirable social values. Such campaigns may also be an opportunity for a social actor to demonstrate moral or cultural legitimacy by asserting prescriptive and proscriptive cultural values and self-deputising itself as an arbiter on such matters, for example, when anti-paedophile organisations assert an ability to protect vulnerable youth. In doing so they may accumulate forms of cultural and symbolic capital, yet economic, political and social capital also shape the reception and legitimacy of such efforts.<sup>24</sup> Such capital may be deliberately solicited, for example, through merchandise sales online, or may be an unintended consequence of mediated exposure.

We may consider how engaging and responding to mediated vigilantism in turn shape conditions of mediated visibility for individuals. Digital vigilantism amounts to a coordinated attack on a targeted individual. The immediate consequences, intended or otherwise, may include damaging the public standing of the individual under scrutiny, as well as those who may share affiliations such as their family, political party, or employer. Regarding the latter, recent cases in the Anglo-American context suggest that job loss is a common tactic, and that employability is a common rhetorical feature when expressing targeted outrage.<sup>25</sup> While an individual target may be centred in such incidents, we must consider how broader audiences are also implicated. The harm invoked is only made possible through the shared understanding that a target is rendered visible to a broader audience, whether that audience is actively watching, or is manifest primarily as an imagined potential. Yet such denunciations play a disciplining role to audience members who may identify with the target, for example, as members of the same minority group. And while the target is the primary recipient of social harm, those participating may also experience unanticipated and unwanted scrutiny, as seen in one of the opening examples. For this reason many individuals and organisations employ a strategic approach to visibility, for example by using 'throwaway' accounts not

linked to any other identifiable information, as well as removing content from the Internet after a limited period of time. While instances of digital vigilantism generate attention as they unfold, scholars remain unaware of longer-term consequences for those who participate or are otherwise implicated in these practices. Individuals who have been exposed to denunciation cope with a compromised reputation, a process shaped both by cultural and technical features. Search-engine manipulation stands as a prominent means to sanitise one's online presence, albeit one that remains unaffordable to many.<sup>26</sup> Among more commonly proposed remedies to coordinated harassment – especially of women – is for the target to close their social media accounts. This amounts to a self-silencing, without addressing the source of harm, nor the possibility of continued abuse.<sup>27</sup> We can imagine that potential remedies will in turn be shaped by public understandings of the risks and outcomes of public shaming. It is equally important to consider the broader viability of mediated denunciation for organisations, as well as the press, government branches and social media platforms, which all may in fact be direct or implicit participants. Digital vigilantism is made up of mediated practices that are currently accessible to virtually any social actor. On basis of the controversy that recent incidents have generated in public discourse, one can anticipate that this may not remain the case in certain jurisdictions over time, either through legal, technical or cultural developments.

## Denunciation and discursive relations

When considering denunciation as a mediated phenomenon, one must consider discursive power relations as they are manifest online and in the press, especially insofar as social actors invoke discourses when seeking various forms of capital. When considering who is permitted to speak about which individuals and which topics in the context of denunciation, several categories of relationships warrant scrutiny. First, the target of denunciation in relation to campaign initiators, who in turn invoke (a) other participants who may share content, add commentary, and even add additional information about target, (b) an often immeasurable audience of media users (social media, but also other media such as the press), and (c) a more abstract understanding of society from which the target – and the broader communities to which they belong – may be expelled. This can remain in abstract as a kind of existential dread, or as specific imagined others such as a future employer. On first pass this appears to be a largely unidirectional relation, with one party acting, others mediating (or mitigating), and another being acted upon, and possibly reacting. In practice we may anticipate a more complex set of communicative relations. Here we may consider if specific categories of social actors are in a more advantageous position to claim what is deemed offensive or otherwise actionable. Conversely, are there conditions where targets are able to refute the denunciation levelled against them? If we accept the premise that such denunciations take place in the closest contemporary equivalent to a public sphere, and that disadvantaged communities have historically been and continue to be poorly served by such platforms, it stands to reason that a reliance on platforms like Twitter to air social grievances may exacerbate categorical forms of discrimination.<sup>28</sup>

Second, we can consider discursive relations between participants and the state. Denunciation may be understood alongside other kinds of user-led forms of justice

seeking that together express a crisis of legitimacy of the state. Mediated vigilante figures like the Guardian Angels, Bernhard Getz and even fictional characters like Paul Kersey (from the *Death Wish* franchise) signify a failure of public order and governance, and thus directly challenge the effectiveness and legitimacy of the state, most notably in terms of a supposed monopolisation of the legitimate force. Although physical violence is a possible outcome in digital vigilantism, we need to reconsider cultural violence as it now relates to harms that can be invoked through prominent representations of the individual. These forms of violence are less explicitly regulated by governments, although we may expect this to change through law, policy and public discourse. Additionally, vigilante groups may challenge states, but are also often a manifestation of hegemonic cultural values (and in some cases reinforce hegemonic economic relations). In response police and other state branches may either ignore, sanction or support vigilante groups. Their official position will likely differ based on characteristics of movements, which becomes difficult if public discourse contributes to a sense of equivalence among various cases of shaming.

Third we can consider how the press contributes to making shaming meaningful. Digital media enable participants and supporters to offer their own commentary that journalists (among others) can either support or contest by integrating it into press coverage of incidents. Beyond journalists, it is important to consider the discursive influence of other public figures such as politicians, academics, public intellectuals, and others who happen to command a sizeable (social media) audience, notably when commenting on controversial or otherwise culturally significant events. As was the case when social media surveillance first entered public consciousness, there is a polyphony of social actors who can collaborate or compete to make mediated denunciation meaningful.

Finally, we may consider social media platforms as passive venues for the above exchanges, yet platform designers shape how digital vigilantism itself is expressed when implementing and commenting on changes made these spaces. As discussed above it is also necessary to consider barriers preventing social actors to access devices and platforms, as well as the necessary forms of literacy and legitimacy.<sup>29</sup> Addressing these so-called digital divides may be regarded as a democratising force, or a levelling of the playing field. Yet this may be a naive view that overlooks reproduction of discrimination not only in its effects, but even in attempts to engage in online denunciations, such that addressing systemic forms of discrimination may remain as difficult as ever. This is evident when considering feminist forms of justice seeking that aim to raise awareness and denounce pervasive forms of gender-based harassment and discrimination online.<sup>30</sup> Familiarity with technology is also important to consider when studying public discourse, as any account of digital vigilantism places demands on its audience. Media venues vary in terms of expectations about knowledge of terminologies, platforms, and prescribed behaviour for the audience-as-digital media user. For example, if addressing a largely non-media-savvy audience, journalists may encounter less resistance when asserting specific understandings of concepts, such as platform-specific parlance.

When studying mediated vigilante campaigns, it often seems that press and other forms of reporting are much more prominent than the campaign itself. A denunciatory post on a message board may generate a few thousand views, but this visibility may be greatly amplified if a journalist covers the incident. We may even question the

distinction between weaponised attempts to make someone visible and press coverage of it, notably if participants are strategically seeking the attention of the press. Thus journalists and others commenting on such incidents may have a bearing not only on how it is rendered meaningful, but may also contribute to harm inflicted on the target, and perhaps a broader community under scrutiny. In further questioning the distinction between platforms that propagate these campaigns, and those that claim to merely report on them, Reddit and the Dutch platform *GeenStijl* are examples of venues where denunciations can occur, but also that can publish reports and editorials on such events. Yet this same confluence can be found in tabloid media, which may either engage in its own denunciations or give coverage to existing campaigns.<sup>31</sup>

### Towards a procedural model of digitally mediated vigilantism

We may consider how digital vigilantism as a set of practices is rendered meaningful in public discourse. Moreover, who is permitted to speak about these practices? Who has 'access to or control over public discourse' on this topic.<sup>32</sup> A first step towards such an analysis is to consider how to conceptualise digital vigilantism, in order to identify concepts that inform and possibly direct media coverage of this content. The tentative model (Figure 1) draws on scholarly literature and empirical instances of shaming in order to consider digital vigilantism as a process consisting of distinct stages, themselves consisting of specific practices.

Past studies of online harassment and surveillance account for the discursive and collective dimensions of mediated shaming. This can be understood as a procedural model structured in terms of policing, discovery and denunciation of offensive content, with each step composed of specific practices. Mediated policing produces conditions

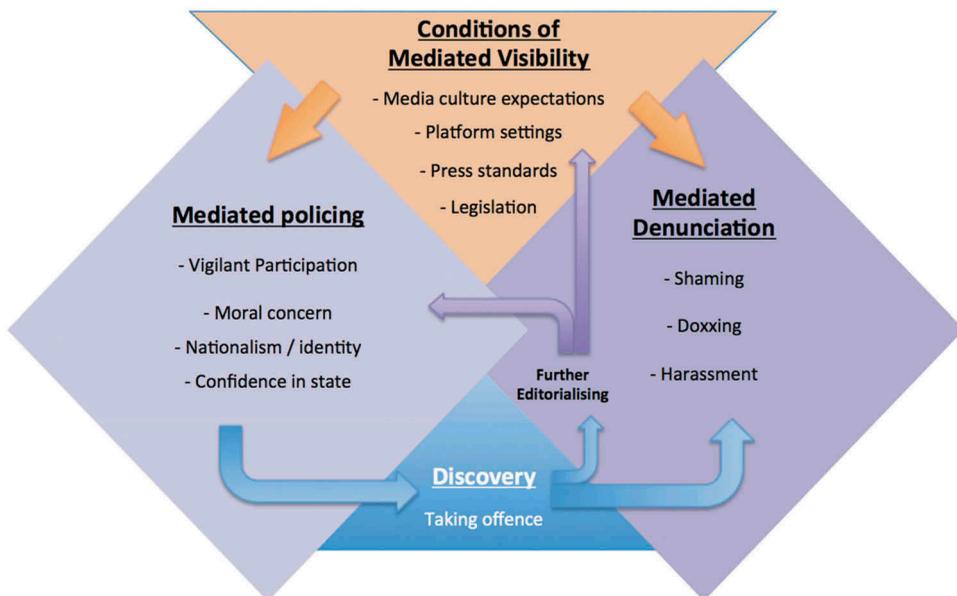


Figure 1. A tentative model of digital vigilantism.

that facilitate the discovery of offences, which in turn trigger denunciatory practices. These are shaped by conditions of visibility, and in turn may re-shape these conditions through public debate, and even legislative and platform-based reforms. Each practice can involve a range of social actors, and ostensibly any public figure may at least attempt to engage in these steps and attract an audience in the process, with or without formal organisational support.

As a set of coordinated activities as well as a way of relating, digital vigilantism is characterised by seeking and making use of targets' personal information, rendering them visible to public scrutiny. The study of these developments can best be addressed by considering **visibility** as a fundamental condition of social life, and one that is being multilaterally renegotiated through digital media use in a range of contexts, including policing as well as moral and civic practices.<sup>33</sup> Alongside embodied visibility, individuals are primarily known to the public through their online presence, coupled with public records and other open sources. We may consider conditions of visibility being shaped by factors including legislation, social media platform terms of use, journalistic practices and standards, and broader cultural expectations about concepts like privacy, publicity and reputation.

**Mediated policing** refers to pre-emptive potential for a campaign, but can also be understood as a more general resting state for digital media users. This denotes an awareness and consideration of steps taken in anticipation or in response to a perceived offence. It is expressed through digital and conventional media through which the audience is implored to adopt a specific predisposition. This can be accomplished by making an explicit appeal to an audience (for example, in the description of a Facebook group, or in a televised program), or by the media actor implicitly prescribing an attitude through their conduct, or in press coverage of events. We can further distinguish between a reactive mode of policing, and deliberately seeking out actionable content, for example, when the Russian group Occupy Paedophilia engages in so-called 'safaris'.<sup>34</sup> Such sentiment may be expressed in press releases and journalistic coverage of campaigns. Likewise, social actors themselves can express their motivations and experiences when engaging in mediated policing as individuals or on behalf of an institution.

As a key dimension of mediated policing, **vigilance** entails a willingness to scrutinise social(ly mediated) life, and to act on offending events if they occur. It is partly informed by a culture of vigilantism that appeals to citizens to take matters into their own hands, but also a mediated culture of participation that solicits more active involvement from media users, albeit under specific circumstances.<sup>35</sup> Vigilance can be understood as steps and measures taken by participating members, including the audience. These are often found in instructions on websites, or implicitly mentioned in newspaper editorials. Vigilance may be framed as passive through listening or lurking in the case of digitally mediated social interactions. Participants do not initially have an active presence, but are rather positioned to catch an offensive act, or recover offensive content. Other forms of activity related to mediated interfaces ('liking'; 'sharing'; 'subscribing'; 'commenting') are both empowering by conferring agency to the audience-as-participant, but also generate information about them that can be retained by group administrators, and platform operators more exclusively. This position in turn sets up the expectation of actions among participants later upon discovery of an offence and target. We can consider a distinction between the participation of primary media actors (witnessing events;

trawling through a profile archive), and of audiences (sharing and reacting to content). Participation-as-vigilance serves to sustain groups in terms of membership, but also to generate diverse forms of capital.<sup>36</sup> More recent forms of measuring capital like the so-called engagement economy are trialled as ways to reproduce economic and social capital, and are expressed as strategies to capitalise on social media platforms.<sup>37</sup> In looking at mediated content distributed by individuals, one can work towards an account of vigilance by considering the kind of resting state that audiences are expected to take (how is their consumption of media content described? How is it related to a broader societal context, including public spaces and specific national or regional territories?). Researchers can also consider the kinds of active steps participants are expected to take in the discovery of offences, for example, by noting references to how citizens are expected to perform when looking out for offences, and whether this performance is prescribed, proscribed or simply anticipated. Furthermore, they may consider how the categorical affiliations and broader social status of those who participate are expressed in reports on such cases.

Mediated policing, as represented in personal blogs, government websites or even reality television, requires an emphasis on specific **moral concerns** that warrant citizen scrutiny and intervention. For example, the feminist blog 'Movethefuckoverbro' addresses the gendered use of space through its description, as well as through discussions in the comments section.<sup>38</sup> Scholars may consider whether the offence is framed as part of a broader social condition, or alternatively as an isolated incident for which the perpetrator is solely responsible. In addition to **proscriptive** elements, media actors may appeal to **prescriptive** characteristics that are deemed to be desirable or necessary to preserve through participation. Recent examples are often territorial, making reference to (the qualities of) a particular country, city or region and/or a related identity or culture that needs to be actively defended. They may also make reference to nostalgic or otherwise idealised social conditions that are endangered by the perceived offence. Research should consider the grounds upon which offences are framed. Along with concrete territories, we can also consider how abstract values such as tradition, equality, progress, justice and even cleanliness are expressed in relation to denunciations.

Digital vigilantism may also reflect a troubled **relation between citizen and state**. Practices that mobilise individuals to identify and denounce other individuals may be expressed as an explicit criticism of the ineffectiveness or unwillingness of the police to intervene in offences, or even the ineligibility of the state to govern over mediated offences. In this sense, campaigns may echo early models of self-policing online where these spaces are seen as independent of any single territory and its governance, a sentiment that is carried into certain strains of cyber-libertarianism and is further complicated through the apparent dissolution of any clear distinction between online and offline activity.<sup>39</sup> It may otherwise simply be an obviation of police duties, for example in the investigation of the Boston Marathon Bombing, or 2011 riot in Vancouver.<sup>40</sup> As such it is worth considering how various branches of the state are described in relation to specific offences, but also in relation to extrajudicial responses. Researchers can consider references to a perceived need to defer to the police or courts, and expressions in favour of 'due process', as well as concepts with judicial purchase such as proportionality or privacy. Relatedly, they may consider references to a lack of

confidence in the police and courts based on a variety of grounds (lack of capabilities; lack of interest; lack of funding; lack of technical expertise).

During the course of mediated policing, participants may come to **discover** offensive conduct, either by witnessing and presumably recording and uploading an embodied offence, or by proactively searching for objectionable content in a target's online presence. They may spread this content through mobile devices onto social platforms, and may add editorial content that serves to reproduce mediated policing. For example it may be expressed to sustain awareness of an issue, or a general state of vigilance, but it also may be expressed more instrumentally through an attempt to sustain the funding and legitimacy of an organisation. If conditions allow, an event can be collectively agreed upon as offensive and actionable, and can lead to response and re-editorialising of context. Yet this implies a certain amount of capital and privilege on the part of those advocating for something deemed offensive, including on behalf of the initiator. An offence may be received as not 'offensive' enough to warrant a substantial response. Likewise, a considerable counter-denunciation may emerge if the initial participant is deemed to have breached social norms. More recent accusations of digital content being manipulated ('photoshopped'), or simply being 'fake news' also offer opportunities to negate the validity of offensive content.

In terms of responses to an offence, we may also distinguish between a first and second order form of moralising, where either the initial offence is denounced, or the response of the individual under scrutiny.<sup>41</sup> To this we may consider a third order offence, in which the response of the participants in a shaming campaign (as described below) are assessed and possibly denounced. Taken together, these may lead to a reconsideration of conditions of visibility that render shaming and vigilance possible. In such responses researchers may consider how visibility as a relatively evasive notion is expressed through coverage of these ordeals. For instance, what aspects of a target or participant's social life are said to be rendered visible, and which individuals, devices or organisations are framed as being responsible for this exposure?

Mediated **denunciation** is sustained by an initial discovery, but also by circulation through digital media by a range of actors. Participants impose denunciatory values and opinions upon a target, who is identified and scrutinised through their personal information, but also through their reputation. Here scholars should be attentive to how the social status of a target is expressed in public discourse, as well as the grounds upon which they are being **shamed** (the offence in question; behaviour more generally; perceptions of their character; categorical affiliation). Shaming may be considered as either reintegrative or stigmatising depending on the tone through which the editorialising against them is expressed, as well as whether there is an expressed possibility or actualised rehabilitation and reintegration of the target.<sup>42</sup>

Likewise, participants engage in **doxing** to render a target visible in a multifaceted and lasting way by gathering and publishing any available information about them and inviting audiences to do the same, as well as to circulate any available information about the target.<sup>43</sup> Scholars should consider the categories of information that participants may seek out (medical; financial; employment; family; location; digital media history), as well as any that may seemingly be excluded from scrutiny, especially in the case of information that may render others visible, such as relatives or members of a shared community. Scholars should also consider how information gathering and distribution

practices are described in, for example, guidelines for groups on social platforms that identify offensive individuals (ex: Reddit group descriptions). Denunciation may provoke other forms of mediated and embodied activities, including **harassment** and bullying, threats, and physical violence, often overlapping with gendered persecution and racism. As for longer-term outcomes, researchers can also consider how the reputation and broader social standing of the target and participants are understood and expressed both in news reports as well as accounts by participants. Here the reputation of the target may both be compromised through denunciatory coverage, but also explicitly reported on in references to the aftermath of the campaign.<sup>44</sup> They may consider references to detrimental life events for targets, for example, an inability to sustain employment, being excommunicated from their community, in addition to physical interventions.

## Discussion

Researchers may take into account a broad range of data that renders mediated vigilante campaigns meaningful. This includes media produced by campaigns themselves, commentary by individual unaffiliated actors, as well as commentary generated by other participants in reaction to initial posts. Likewise, journalistic coverage often features prominently when searching cases, including platforms that straddle a boundary between reporting the news and utilising social media to propagate campaigns. Still other social actors may seek to comment on these cases and phenomena more broadly through guides and reports. In the context of gender-based harassment, online resources<sup>45</sup> provide direction for those on the receiving end of vitriol, and in turn shape public perception of such incidents.

Digitally mediated vigilantism constitutes a troubling set of practices that combines the scrutiny of targeted individuals with denunciation and a range of social harms. As such, researchers should remain attentive that even supposedly user-led activities inform and in turn are informed by other institutions. The press have historically engaged in citizen shaming, which begins to overlap with digital media practices as content from these venues attract digital media users. This amounts to a kind of hybrid media system that not only involves individual user input, but may also shape relations between individuals, states, and organisations.<sup>46</sup> While journalism more broadly involves rendering people visible through (as one example) crime reporting, the concern with digital vigilantism is the combination of a denunciatory tone with the possibility of audiences further participating in retaliatory activities. Contemporary denunciations amount to a proliferation of editorialising, coupled with volatile and harmful data handling practices. The press are composed of broadsheet and tabloid newspapers with an affiliated digital media presence such as *The Daily Mail*, but also media agencies that have been established and operate entirely online, such as *GeenStijl.nl*, *Breitbart* and *TheRebel.ca*. Other organisations may also participate in mediated shaming. These include citizen-led initiatives, civil society organisations, as well as profit-oriented groups. Such groups may cultivate their own mediated visibility through branding strategies on social media. They often leverage audience cynicism in order to capitalise on the visibility and denunciation of targeted individuals.<sup>47</sup> As well, police and other state agencies adopt diverging strategies towards mediated shaming. In some instances

they advocate for and enforce privacy and data protection laws as legal mechanisms against coordinated attacks on an individual's social standing. Campaigns may cross legal thresholds if threats are uttered, and may warrant police intervention. Yet in a global context states are also making use of mediated denunciations as a potential form of 'soft power', for example, in response to unpaid taxes.<sup>48</sup> It is not always clear if citizens are self-mobilised, or operating with the implicit deputisation by the state. Finally, digital vigilantism occurs through social media platforms such as Twitter, Facebook, Reddit and VKontakte along with other telecommunication and digital media services. These services provide specific affordances for the above social actors to engage in shaming as well as generate various forms of capital. And while they may take measures to minimise harmful outcomes such as doxing and harassment, their business models benefit from viral spikes in online engagement following controversial shaming campaigns, to the extent that they may tolerate or even cultivate such coordinated forms of social harm.

## Notes

1. Scott, "In Defense of Call-out Culture."
2. Trottier, "Digital vigilantism."
3. Treem et al., "Social media use in organizations"; and van Dijck, et al., "Understanding Social Media Logic."
4. Nhan et al., "Digitalantism"; Schneider et al., "Social Media and the 2011 Vancouver Riot."; Jane, "Online misogyny and feminist digilantism."
5. Johnston, "What is vigilantism?"
6. Moncada, "Varieties of vigilantism."
7. See note 5 above.
8. Vitis et al., "Dick pics on blast."
9. Blitz, "Lies, Line Drawing, and (Deep) Fake News."
10. Haggard, "North Korea and the Sony hack"; Mansfield-Devine, "The Ashley Madison affair"; Persily "The 2016 U.S. Election"; and Ferrara "Disinformation and social bot operations."
11. See <https://abc30.com/news/social-media-users-help-id-philadelphia-gay-bashing-suspects/312333/>.
12. Kohm, "Naming, shaming and criminal justice."
13. Trottier, *Social Media as Surveillance*.
14. Suler, "The Online Disinhibition Effect."
15. Pearce et al., "Performing honor online."
16. Hedenus et al., "Explaining the Data Double."
17. Johnston, "What is vigilantism?", 222.
18. van Dijck et al., "Understanding Social Media Logic."
19. Skoric et al., "Online shaming in the Asian context."
20. c.f. Trottier, *Social Media as Surveillance*.
21. Braithwaite, "It's about ethics."
22. Mendes et al., "# MeToo and the promise and pitfalls."
23. Witteborn et al., "Diaosi as a way of relating."
24. Dupont, "Security in the age of networks"; and Myles et al., "Leveraging Visibility."
25. Di Placido, "The Strange Story of James Gunn"; and See <https://www.facebook.com/shaunking/posts/1145448458827404>.
26. Ronson, *So You've Been Publicly Shamed*.
27. Lenhart et al., "Online harassment."
28. Papacharissi, "The virtual sphere"; and Fuchs, "Social media and the public sphere."

29. van Deursen et al., "Internet skills and the digital divide"; Harambam et al., "The contentious gap"; and Goode, "The digital identity divide."
30. Vitis et al., "Dick pics on blast"; Jane, "Online misogyny and feminist digilantism"; and Lenhart et al., *Online harassment*.
31. Meyer, "Evil monsters and cunning perverts"; and Campbell "Policing paedophilia."
32. van Dijk, "Critical Discourse Analysis."
33. Brighenti, "Visibility"; Thompson, "The New Visibility"; Goldsmith, "Policing's new visibility"; and Trottier, "Scandal mining."
34. Favarel-Garrigues et al., "Violer la loi pour maintenir l'ordre."
35. Abrahams, *Vigilant Citizens*; and Barney et al., *The Participatory Condition*.
36. Dupont, "Security in the age of networks."
37. McGonigal, *Engagement Economy*.
38. Rose, "Meet the Orthodox Woman."
39. Wall et al., "Policing Diversity in the Digital Age"; Dahlberg, "Cyber-libertarianism 2.0."
40. Nhan et al., "Digilantism"; and Schneider et al., "Social Media and the 2011 Vancouver Riot."
41. c.f. Thompson, "The New Visibility."
42. Braithwaite, *Crime, shame and reintegration*.
43. Douglas, "Doxing: a conceptual analysis."
44. Solove, *The Future of Reputation*.
45. See <http://blog.unburntwitch.com/post/99694581114/what-to-expect-when-youre-expecting-the-internet> and <https://arstechnica.com/security/2015/03/anti-doxing-strategy-or-how-to-avoid-50-qurans-and-287-of-chick-fil-a/>.
46. Chadwick, "The political information cycle."
47. c.f. Fieschi et al., "Trust, cynicism and populist anti-politics."
48. See note 19 above.

## Disclosure statement

No potential conflict of interest was reported by the author.

## Funding

This work was supported by the Netherlands Organisation for Scientific Research (NWO) [276-45-004];

## Notes on contributor

*Daniel Trottier* is an Associate Professor at the Department of Media and Communication of Erasmus University Rotterdam. His current research considers the use of digital media for the purposes of scrutiny, denunciation and shaming. Daniel completed a PhD in Sociology at Queen's University, Canada.

## Bibliography

- Abrahams, R. G. *Vigilant Citizens: Vigilantism and the State*. Cambridge: Polity Press, 1998.
- Barney, D., G. Coleman, C. Ross, J. Sterne, and T. Tembeck, eds. *The Participatory Condition in the Digital Age*. Minneapolis: University of Minnesota Press, 2016.
- Blitz, M. J. "Lies, Line Drawing, and (Deep) Fake News." *Oklahoma Law Review* 71, no. 1 (2018): 59–116.
- Braithwaite, A. "It's about Ethics in Games Journalism? Gamergaters and Geek Masculinity." *Social Media+ Society* 2, no. 4 (2016): 1–10. doi:10.1177/2056305116672484.
- Braithwaite, J. *Crime, Shame and Reintegration*. Cambridge: Cambridge University Press, 1989.

- Brighenti, A. "Visibility: A Category for the Social Sciences." *Current Sociology* 55, no. 3 (2007): 323–342. doi:10.1177/0011392107076079.
- Campbell, E. "Policing Paedophilia: Assembling Bodies, Spaces and Things." *Crime, Media, Culture* 12, no. 3 (2016): 345–365. doi:10.1177/1741659015623598.
- Chadwick, A. "The Political Information Cycle in a Hybrid News System: The British Prime Minister and the "Bullygate" Affair." *The International Journal of Press/Politics* 16, no. 1 (2011): 3–29. doi:10.1177/1940161210384730.
- Dahlberg, L. "Cyber-Libertarianism 2.0: A Discourse Theory/Critical Political Economy Examination." *Cultural Politics* 6, no. 3 (2010): 331–356. doi:10.2752/175174310X12750685679753.
- Di Placido, D. "The Strange Story of James Gunn." *Forbes*. Accessed July 21, 2018. <https://www.forbes.com/sites/danidiplacido/2018/07/21/the-strange-story-of-james-gunn>
- Douglas, D. "Doxing: A Conceptual Analysis." *Ethics and Information Technology* 18, no. 3 (2016): 199–210. doi:10.1007/s10676-016-9406-0.
- Dugan, M. J. "Megan's Law or Sarah's Law-A Comparative Analysis of Public Notification Statutes in the United States and England." *Loyola of Los Angeles International and Comparative Law Review* 23 (2001): 617.
- Dupont, B. "Security in the Age of Networks." *Policing and Society* 14, no. 1 (2004): 76–91. doi:10.1080/1043946042000181575.
- Ferrara, E. "Disinformation and Social Bot Operations in the Run up to the 2017 French Presidential Election." *First Monday* 22, no. 8 (2017). doi:10.5210/fm.v22i18.8005.
- Fieschi, C., and P. Heywood. "Trust, Cynicism and Populist Anti-Politics." *Journal of Political Ideologies* 9, no. 3 (2004): 289–309. doi:10.1080/1356931042000263537.
- Fuchs, C. "Social Media and the Public Sphere." *TripleC: Open Access Journal for a Global Sustainable Information Society* 12, no. 1 (2015): 57–101. doi:10.31269/triplec.v12i1.552.
- Goldsmith, A. J. "Policing's New Visibility." *The British Journal of Criminology* 50, no. 5 (2010): 914–934. doi:10.1093/bjc/azq033.
- Goode, J. "The Digital Identity Divide: How Technology Knowledge Impacts College Students." *New Media & Society* 12, no. 3 (2010): 497–513. doi:10.1177/1461444809343560.
- Haggard, S., and J. R. Lindsay. "North Korea and the Sony Hack: Exporting Instability through Cyberspace." *AsiaPacific Issues* 117 (2015): 1–8.
- Harambam, J., S. Aupers, and D. Houtman. "The Contentious Gap: From Digital Divide to Cultural Beliefs about Online Interactions." *Information, Communication & Society* 16, no. 7 (2013): 1093–1114. doi:10.1080/1369118X.2012.687006.
- Hedenus, A., and C. Backman. "Explaining the Data Double: Confessions and Self-Examinations in Job Recruitments." *Surveillance & Society* 15, no. 5 (2017): 640–654. doi:10.24908/ss.v15i5.6380.
- Jane, E. A. "Online Misogyny and Feminist Digilantism." *Continuum* 30, no. 3 (2016): 284–297. doi:10.1080/10304312.2016.1166560.
- Johnston, L. "What Is Vigilantism?" *The British Journal of Criminology* 36, no. 2 (1996): 220–236. doi:10.1093/oxfordjournals.bjc.a014083.
- Kohm, S. A. "Naming, Shaming and Criminal Justice: Mass-Mediated Humiliation as Entertainment and Punishment." *Crime, Media, Culture* 5, no. 2 (2009): 188–205. doi:10.1177/1741659009335724.
- Lenhart, A., M. Ybarra, K. Zickuhr, and M. Price-Feeney. *Online Harassment, Digital Abuse, and Cyberstalking in America*. Data and Society Research Institute, New York, NY, 2016.
- Mae, K. "We with the Pitchforks." *Abandoning Pretense*. Accessed June 4, 2016. <http://abandoningpretense.com/2016/06/we-with-the-pitchforks.html>
- Mansfield-Devine, S. "The Ashley Madison Affair." *Network Society* 9 (2017): 8–16. doi:10.1016/S1353-4858(15)30080-5.
- McGonigal, J. *Engagement Economy: The Future of Massively Scaled Collaboration and Participation*. Palo Alto, CA: Institute for the Future, 2008.
- Mendes, K., J. Ringrose, and J. Keller. "# MeToo and the Promise and Pitfalls of Challenging Rape Culture through Digital Feminist Activism." *European Journal of Women's Studies* 25, no. 2 (2018): 236–246. doi:10.1177/1350506818765318.

- Myles, D., and D. Trottier. "Leveraging Visibility, Gaining Capital? Social Media Use in the Fight Against Child Abusers: The Case of The Judge Beauce." *Social Media + Society* 3, no. 1: doi: [10.1177/2056305117691998](https://doi.org/10.1177/2056305117691998).
- Meyer, A. "Evil Monsters and Cunning Perverts: Representing and Regulating the Dangerous Paedophile." In *Popular Culture, Crime and Social Control* edited by Mathieu Deflem, 195–217. Bingley: Emerald Group Publishing Limited, 2010.
- Moncada, E. "Varieties of Vigilantism: Conceptual Discord, Meaning and Strategies." *Global Crime* 18, no. 4 (2017): 403–423. doi:[10.1080/17440572.2017.1374183](https://doi.org/10.1080/17440572.2017.1374183).
- Nhan, J., L. Huey, and R. Broll. "Digilantism: An Analysis of Crowdsourcing and the Boston Marathon Bombings." *The British Journal of Criminology* 57, no. 2 (2017): 341–361.
- Papacharissi, Z. "The Virtual Sphere: The Internet as a Public Sphere." *New Media & Society* 4, no. 1 (2002): 9–27. doi:[10.1177/1461444022226244](https://doi.org/10.1177/1461444022226244).
- Pearce, K. E., and J. Vitak. "Performing Honor Online: The Affordances of Social Media for Surveillance and Impression Management in an Honor Culture." *New Media & Society* 18, no. 11 (2016): 2595–2612. doi:[10.1177/1461444815600279](https://doi.org/10.1177/1461444815600279).
- Persily, N. "The 2016 U.S. Election: Can Democracy Survive the Internet?" *Journal of Democracy* 28, no. 2 (2017): 63–76. doi:[10.1353/jod.2017.0025](https://doi.org/10.1353/jod.2017.0025).
- Ronson, J. *So You've Been Publicly Shamed*. New York: Riverhead Books, 2015.
- Rose, R. "Meet the Orthodox Woman behind 'Move the Fuck Over, Bro' Blog." Jezebel. Accessed January 15, 2014. <https://jezebel.com/meet-the-orthodox-woman-behind-move-the-fuck-over-bro-1501688008>
- Schneider, C. J., and D. Trottier, "Social Media and the 2011 Vancouver Riot", in 40th Anniversary of Studies in Symbolic Interaction (Studies in Symbolic Interaction, Volume 40), edited by N. K. Denzin, 335–362. Bingley: Emerald Group Publishing, 2013.
- Scott, S. "In Defense of Call-Out Culture." CityArts. Accessed February 1, 2018. [www.cityartsmagazine.com/defense-call-culture/](http://www.cityartsmagazine.com/defense-call-culture/)
- Skoric, M. M., J. P. E. Chua, M. A. Liew, K. H. Wong, and P. J. Yeo. "Online Shaming in the Asian Context: Community Empowerment or Civic Vigilantism?" *Surveillance & Society* 8, no. 2 (2010): 181–199. doi:[10.24908/ss.v8i2.3485](https://doi.org/10.24908/ss.v8i2.3485).
- Solove, D. *The Future of Reputation*. New Haven, CT: Yale University Press, 2007.
- Suler, J. "The Online Disinhibition Effect." *CyberPsychology & Behaviour* 7, no. 3 (2004): 321–326. doi:[10.1089/1094931041291295](https://doi.org/10.1089/1094931041291295).
- Thompson, J. B. "The New Visibility." *Theory, Culture & Society* 22, no. 6 (2005): 31–51. doi:[10.1177/0263276405059413](https://doi.org/10.1177/0263276405059413).
- Treem, J. W., and P. M. Leonardi. "Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association." *Annals of the International Communication Association* 36, no. 1 (2013): 143–189. doi:[10.1080/23808985.2013.11679130](https://doi.org/10.1080/23808985.2013.11679130).
- Trottier, D. *Social Media as Surveillance: Rethinking Visibility in a Converging World*. Farnham: Ashgate, 2012.
- Trottier, D. "Digital vigilantism as weaponisation of visibility." *Philosophy & Technology* 30, no. 1 (2017): 55–72. doi:[10.1007/s13347-016-0216-4](https://doi.org/10.1007/s13347-016-0216-4).
- Trottier, D. "Scandal mining: political nobodies and remediated visibility." *Media Culture and Society* 40, no. 6 (2018): 893–908. doi:[10.1177/0163443717734408](https://doi.org/10.1177/0163443717734408).
- van Deursen, A., and J. van Dijk. "Internet Skills and the Digital Divide." *New Media & Society* 13, no. 6 (2011): 893–911. doi:[10.1177/1461444810386774](https://doi.org/10.1177/1461444810386774).
- van Dijck, J., and T. Poell. "Understanding Social Media Logic." *Media and Communication* 1, no. 1 (2013): 2–14. doi:[10.17645/mac.v1i1.70](https://doi.org/10.17645/mac.v1i1.70).
- van Dijk, T. A. "Critical Discourse Analysis." In *The Handbook of Discourse Analysis, Volume 1*, edited by D. Annen, H. E. Hamilton, and D. Schiffrin, 466–485. Chichester: Wiley, 2015.

- Vitis, L., and F. Gilmour. "Dick Pics on Blast: A Woman's Resistance to Online Sexual Harassment Using Humour, Art and Instagram." *Crime, Media, Culture* 13, no. 3 (2017): 335–355. doi:[10.1177/1741659016652445](https://doi.org/10.1177/1741659016652445).
- Wall, D. S., and W. Matthew. "Policing Diversity in the Digital Age: Maintaining Order in Virtual Communities." *Criminology and Criminal Justice* 7 (2007): 391–415. doi:[10.1177/1748895807082064](https://doi.org/10.1177/1748895807082064).
- Witteborn, S., and Q. Huang. "Diaosi as a Way of Relating in Contemporary China." In *The Handbook of Communication in Cross-Cultural Perspective*, edited by D. Carbaugh, 142–152. New York: Routledge, 2017.