



The Indian Journal for Research in Law and Management

Open Access Law Journal – Copyright © 2026

Editor-in-Chief – Dr. Muktai Deb Chavan; Publisher – Alden Vas; ISSN: 2583-9896

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International (CC-BY-NC-SA 4.0) License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

TRIAL BY SOCIAL MEDIA: THE SOCIOLOGY OF DIGITAL PUNISHMENT BEFORE THE COURTROOM

~ Rudrali Deshpande

ABSTRACT

The paper analyses the phenomenon of “trial by social media,” differentiating it from traditional trial by media as both a sociological and legal concept. With reference to theories of the informal control of behaviour, stigma, and digital vigilantism, the text points out how different kinds of platforms increase the message of moral indignation through their algorithms, flatten different contexts, and disseminate responsibility among anonymous users. A comparative analysis shows different approaches to regulation in the USA, India, and the EU, particularly looking at the role of Section 230, the Shreya Singhal case, and the Digital Services Act for the understanding of legal liability of platforms. Evidence from Pew Research studies indicates the extent and severity of online harassment, while studies of virality and networked misogyny underline the higher degree of vulnerability of some groups. The author of the paper emphasises that while trials by social media are resistant to established approaches to regulation due to the absence of due process, they nevertheless reveal the failures of institutions and require new forms of accountability. Ultimately, the analysis places digital punishment in larger discussions about justice, governance, and outrage politics. It concludes that both co-regulatory and restorative methods are effective but come with their own challenges relating to censorship and legitimacy.

Keywords: Trial by Social Media, Algorithmic Amplification, Digital Vigilantism, Platform Liability and Politics of Outrage.

1. INTRODUCTION

It is essential to understand that long before social media emerged, it was recognised by sociologists that formal legal institutions were never society's sole mechanism for enforcing norms. The informal factors of social control, such as ostracism, public shaming and communal policing, have acted as a moral police and controlled human behaviour in communities for millennia. The major change in the twenty-first century is not the impulse to judge and punish the transgressors outside the courtroom, but rather the scale, speed and permanence with which that impulse can now be enacted. A single tweet, post, video clip or screenshot can travel to an audience of millions within minutes. This has transformed an individual's private conduct, opinion or decontextualised moment into a matter of global public judgement. This phenomenon is commonly described as "trial by social media". It is a process by which social media platforms such as X(formerly Twitter), Instagram, Facebook and Reddit have informally assumed the role of courts, where often anonymous accounts become prosecutors, the aggregate of commenting and sharing users acts as juries and lastly, harassment campaigns, loss of jobs, cancel culture and reputational destruction have become sentences or punishments. All are typically rendered before or entirely independent of any formal judicial proceeding.

This paper aims to approach trial by social media as a sociological and legal object rather than a journalistic one. It primarily asks three interrelated questions. First, what is it about trials conducted on social media that makes them different from the older forms of trial by the media that relied on old avenues such as newspapers and television? Second, what are the socio-structural and platform-specific explanations for the emergence of such punishments and their spread and subsidence? Third, what are the implications of this phenomenon for people, notions of justice, and social order in more general terms, as well as whether and which corrections are being made?

The article is divided into five segments. In Section 2, it examines the conceptual framework that has been established in communication and media sociology, especially the notion of "digital vigilantism." In Section 3, it contextualises trial by social media in relation to the concept of "trial by media," which prevailed during the television and newspaper period. Section 4 explores mechanisms like technical, psychological, and social factors that produce digital punishment. Section 5 discusses implications and criticisms.

The paper concludes with ideas about what should be done and what research is necessary in the future.

This paper is a qualitative study based on a review of existing literature. The methodology constitutes a synthesis of existing literature and secondary data; it employs established sociology and communication theories such as Durkheim on collective punishment, Goffman's theory of stigma and Trottier's theory of digital vigilantism as a framework and complements it with findings from peer-reviewed research papers like Brady et al.'s PNAS paper on moral-emotional virality, the work done by Marwick & boyd related to context collapse and Suler's theory of online disinhibition and refers to nationally representative statistics from external sources like survey data from the Pew Research Centre. Therefore, in terms of research methodology, this paper can be categorised as a conceptual/theoretical review paper because it develops an argument based on the analysis and research of various scholars rather than using primary empirical methods such as interviews, content analysis of social media posts or experiments.

2. THEORETICAL FRAMEWORK: FROM INFORMAL SOCIAL CONTROL TO DIGITAL VIGILANTISM

According to Émile Durkheim, punishment is an act of restoring the “collective conscience”, which means that punishment has an expressive function as it emphasises what acceptable and unacceptable human behaviour is.¹ Goffman's research on stigma suggests the idea that platforms can stigmatise people and turn one act into a lifetime identity. Thus, these systems describe how shaming works in modern society.² Daniel Trottier's notion of “digital vigilantism” is one of the most influential contemporary theories whereby citizens who feel collectively offended tend to coordinate retaliation through mobile and online tools after feeling offended in a group, regardless of the severity of the wrongdoing from a small breach of the rules to acts of terrorism.³ Trottier expresses this in the form of the “weaponisation of visibility,” meaning that instead of the police being involved in the process, this method functions as an independent means of justice provided by users themselves. The punishment is not simply reflecting on the violation but planning

¹ Durkheim, É., *The Division of Labour in Society* (1893; W.D. Halls trans., Free Press 1984).

² Goffman, E., *Stigma: Notes on the Management of Spoiled Identity* (Prentice-Hall 1963).

³ Trottier, D., *Digital Vigilantism as Weaponisation of Visibility*, 30 *Phil. & Tech.* 55 (2017).

the disclosure of private evidence, such as revealing someone's personal information, professional data or even place of living to implement retaliation.⁴

Trottier's later research differentiates various levels of criminal offence, as well as organised crime and organised condemnation between individuals. The whole process is condensed into a single event taking place in a matter of a few hours, instead of going through several stages, such as investigation, indictment and verdict under legal systems. It is important to note that this process is not carried out according to the laws and rules of conduct in this regard. Social media trials cannot be classified either as a beneficial process or as a negative event and the definite interpretation depends on the position of the target.⁵

Therefore, the legal grounding lies in what Trottier explains sociologically: that public punishment instigated by users through networks that are parallel to official institutions has a legal equivalent. The present state relies on what Weber termed the monopoly on legal violence, implemented in punishment through the presumption of innocence, trials in front of unbiased courts and the right to appeal. These guarantees differentiate state power of punishment from personal revenge. Digital vigilantism, functioning via private, decentralised, and often underground cooperation, does not fall under this framework; nobody can tell who is entitled to declare somebody guilty of a social crime on the Internet, based on what evidence, or determine the appropriate punishment.

3. EVOLUTION OF MEDIA TRIAL TO SOCIAL MEDIA TRIAL

The term “trial by media” predates social media and explains how coverage by the press plays an important role in shaping the public opinion regarding the guilt of the accused before or during the trial, which potentially prejudices juries. Social media has made this more intense and complex by letting elements of the audience participate in the process of trial judgment and exerting invisible influence over the prosecutors and lawyers.

The legal system did not leave trial by media unregulated. India's Contempt of Court Act, 1971 empowers the courts to punish those who publish any information that can prejudice the ongoing proceedings. The Supreme Court, in the case of *Sahara India Real Estate Corp.*

⁴ Ibid.

⁵ Trottier, D., Denunciation and Doxing: Towards a Conceptual Model of Digital Vigilantism, 21 Global Crime 196 (2020).

*Ltd. v. SEBI*⁶ elaborated upon the “doctrine of postponement”, which grants the courts the power to temporarily impose restrictions on prejudicial materials while also respecting the free-press guarantees under Article 19(1)(a) and fair trial rights under Article 21.⁷ England's Contempt of Court Act 1981 lays down that a comparable strict liability rule would apply once the proceedings have commenced. In contrast, the U.S. Supreme Court adopted a much more speech-protective stance in the case of *Nebraska Press Association v. Stuart*,⁸ where it held that prior restraint on trial coverage is unconstitutional and that the right to a fair trial can be protected through in-trial processes rather than prior restraints.

Each doctrine points towards a particular publisher and a specific, duly noted proceeding. Trial by social media, however, destroys both principles. In this situation, responsibility falls on countless anonymous users, and, as a rule, no legal proceedings occur, since most targets of trial are never charged with any crime.

Empirically, Pew Research Centre's survey conducted in 2020 suggests that three-quarters of victims of harassment reported their last incident taking place on social media.

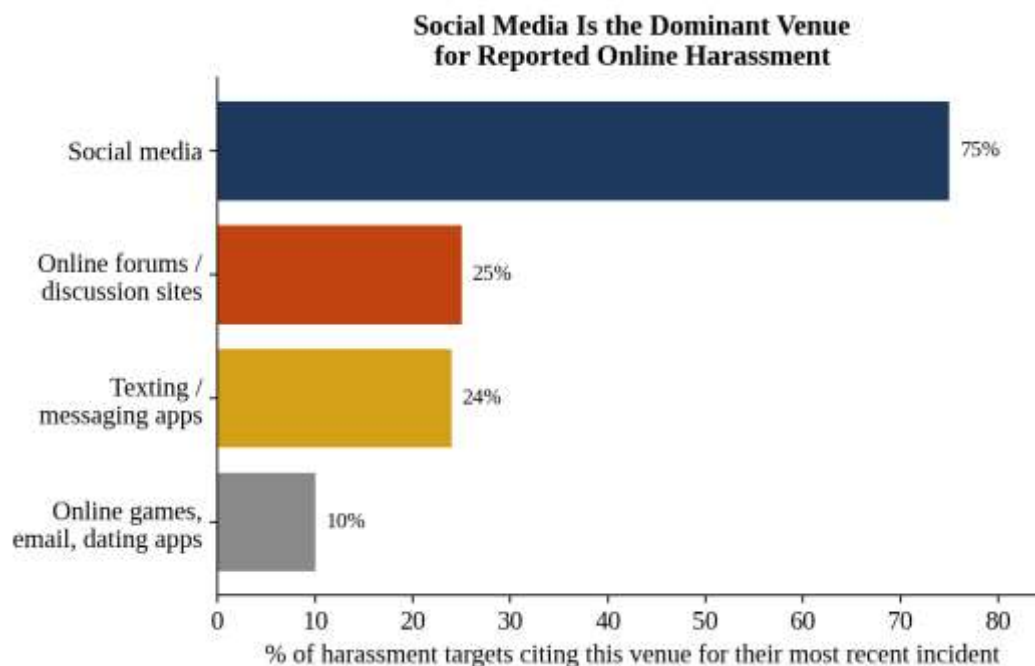


Figure 1: Social media is the dominant venue for online harassment⁹

⁶Sahara India Real Estate Corp. Ltd. v. SEBI, (2012) 10 SCC 603

⁷ibid.

⁸Nebraska Press Association v. Stuart, 427 U.S. 539 (1976)

⁹Pew Research Centre, “The State of Online Harassment” (2021). Nationally representative survey of 10,093 U.S. adults, Sept. 8-13, 2020. Respondents could name multiple venues; figures do not sum to 100%.

The current era features several distinctions from the trial-by-media concept familiar from the broadcast age. One such distinction is the transition from a passive reception of media to active partnership in content creation, whereby users take on the roles of the investigator, the prosecutor and the juror. The socio-legal scholarship exploring this transition refers to the phenomenon of trial by social media, which can be defined as formulating judgements before any court has issued a verdict, thereby often infringing upon the presumption of innocence and the right to privacy. Second, the temporal compression in the current age is much more significant, as the formation of public outrage can take place within minutes due to engaging algorithms as opposed to focusing on the truth. Third and most importantly, the judgment here does not have to concern a person accused of a crime, as it can target anyone who said something inappropriate or made a gesture that was misinterpreted, even if no crime had taken place.

This is the reason why sociologists view trial by social media differently from merely being an advancement of technology. The research about the decline of the presumption of innocence refers to informal penalties exercised against even those who will be cleared later on as "additional punishment" aside from the legal process, whereas criminology here understands that it falls under the more extensive phenomenon of "politics of outrage," where all forms of policing and prosecution happen under networked scrutiny.

4. MECHANISMS OF DIGITAL PUNISHMENT

4.1 PLATFORM ARCHITECTURE AND ALGORITHMIC AMPLIFICATION

The algorithms for recommendations are designed in such a way that engagement is maximised and moral outrage effectively promotes it. As a result, the content that is designed to evoke anger reaches more users than purely neutral content, which results in a feedback loop through which "outrage" becomes visible irrespective of its correctness. In the study published in PNAS by Brady et al. (2017),¹⁰ which analysed about 563,000 tweets, it became apparent that neither pure moral nor pure emotional language was sufficient to attract the increased number of retweets. In contrast, the use

¹⁰ Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralised content in social networks. *Proceedings of the National Academy of Sciences*, 114(28), 7313–7318. Based on an analysis of approximately 563,000 tweets on gun control, same-sex marriage, and climate change.

of both types of language, which is termed "moral-emotional," was found to lead to an increase in retweets by about 20% per word.

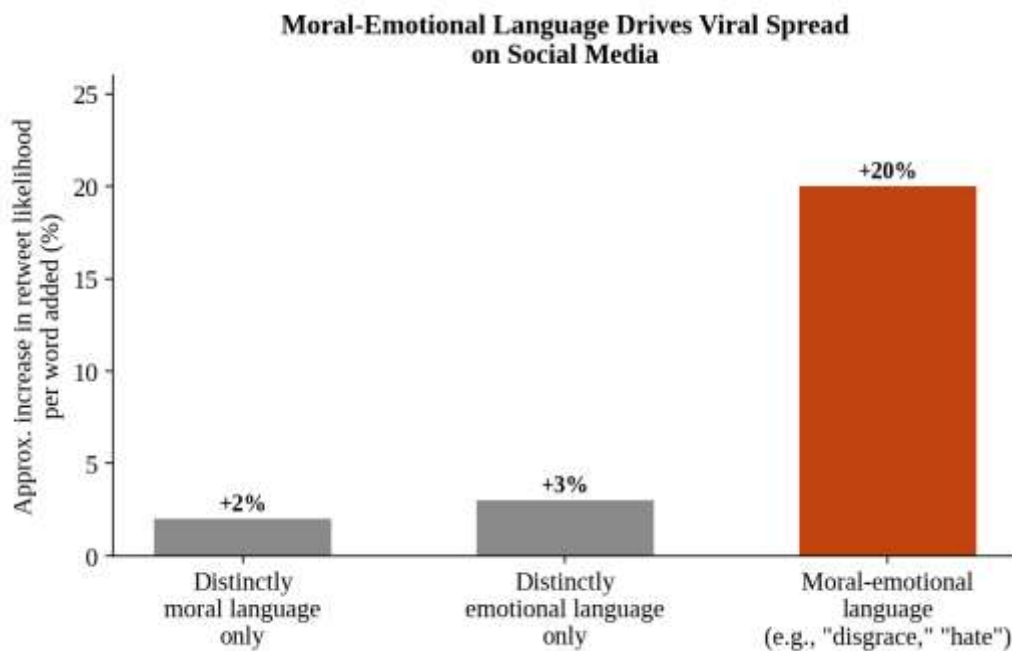


Figure 2: Moral-emotional language drives viral spread¹¹

This means that the information ecosystem wherein the punishment campaigns occur is at odds with the truth of the respective events.

4.2 CONTEXT COLLAPSE AND DECONTEXTUALISATION

Marwick and boyd (2011) explain that Internet platforms have different "imagined audiences," because different groups of people can access the same communication and messages. Hence, some comments made on social media platforms, no matter whether they were made jokingly or a decade ago, can appear before an audience who does not know the situation. In trial by social media campaigns, such images or quotations are often used to depict the person under accusation without any proof.

4.3 ANONYMITY, DISINHIBITION, AND RESPONSIBILITY DIFFUSION

According to Suler's (2004) explanation of the "online disinhibition effect," there are many factors that help explain why people behave in the online space in ways they wouldn't behave when meeting the same person in the real world, namely dissociation from offline identity and diminished accountability due to not being physically visible.

¹¹ Ibid.

When the harm is spread over millions of small contributions from different people, none of them feels responsible for this overall damage. According to Jon Ronson, ordinary people undergo various forms of harassment after one inappropriate comment goes viral, citing, for instance, the case of a charity worker whose life was ruined because of one old photo that resurfaced after many years and put him into a state of prolonged depression and withdrawal from public life.

4.4 LEGAL GROUNDING: PLATFORM LIABILITY LAW AS A STRUCTURAL MECHANISM

The above-mentioned mechanisms function in a particular legal incentive arrangement. In the United States, Section 230 of the Communications Decency Act¹² gives platforms immunity from liability for carrying or even algorithmically amplifying user-generated content that they did not create. Therefore, the need to intervene at the beginning of an emergent pile-on is removed. India arrives at the same conclusion with a different rationale. Section 79 of the IT Act, 2000¹³ provides for a conditional safe harbour, while, in the case of *Shreya Singhal v. Union of India*¹⁴, the Supreme Court invalidated vague Section 66A and interpreted Section 79 in such a way that intermediaries would only be compelled to act if ordered by the court or notified by the government, but not due to private complaints. The IT Rules 2021 and the Grievance Appellate Committee create due-diligence obligations for large platforms, although, following the ruling in *Shreya Singhal*, platforms are still not required to proactively safeguard against the emergence of a pile-on unless instructed to do so.

The European Union (EU) has progressed the furthest in the direction of affirmative obligations. The Digital Services Act (Regulation (EU) 2022/2065)¹⁵, which has been in effect since February 2024, mandates that “Very Large Online Platform Operators” perform systemic risk assessments focused on potential harm magnification and take corresponding mitigation actions, with oversight by the European Commission and national coordinators. Hence, there arises an almost linear continuum of liability: broad

¹² Communications Decency Act (47 U.S.C. § 230)

¹³ Information Technology Act, 2000, § 79 (India).

¹⁴ *Shreya Singhal v. Union of India*, AIR 2015 SC 1523.

¹⁵ Daphne Keller, *Who Do You Sue? State and Platform Hybrid Power over Online Speech*, Hoover Institution Aegis Series Paper No. 2019-01.

immunity in the U.S., conditional immunity in India, and an obligation to mitigate risks arising in the EU.

5. CONSEQUENCES AND CRITIQUES

5.1 DISPROPORTIONALITY AND THE ABSENCE OF PROCEDURAL SAFEGUARDS

Formal justice involves procedural protections such as the evidentiary rules, rights of defence or appeal because the verdict of the state is backed by coercion. Social media trials lack these safeguards, but their consequences can have a serious effect. The reasoning behind this is explained in legal theory as that of “punishment without law,” existing in a legal void and not subject to reform after its original narrative is disproven.

5.2 REPRODUCTION OF INEQUALITY

Rather than being a great equaliser, digital punishment often reinforces inequalities of race, gender, class and nationality. Citron's (2014) study on gender-mediated cyber-harassment states that coordinated attacks are predominantly directed against women, using a combination of reputation attacks with threats of rape and non-consensual revealing of intimate images in a manner that does not happen in the case of punishment for men.¹⁶ This is what feminist scholars call networked misogyny in the name of accountability.

5.3 PSYCHOLOGICAL AND SOCIAL COSTS

People affected experience long-lasting psychological distress and negative effects on their reputation that remain online forever, well beyond any legal repercussions that they may suffer. The Pew Research data show that while the total prevalence level of online harassment accelerated by around 41% since 2017, its severity increased significantly: the level of severe harassment (comprising threats, stalking, and continued harassment) rose from 15% in 2014 to 25% in 2020.¹⁷

¹⁶Citron, D. K., *Hate Crimes in Cyberspace* (Harv. U. Press 2014).

¹⁷Pew Rsch. Ctr., *The State of Online Harassment* (Jan. 13, 2021).

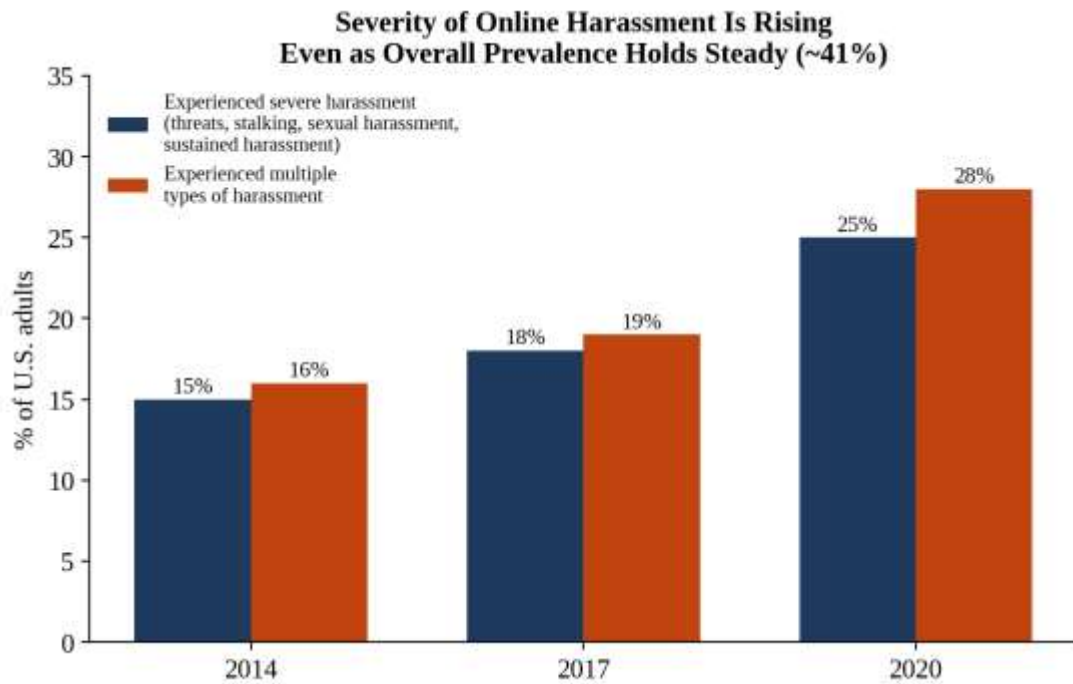


Figure 3: Severity of online harassment is rising¹⁸

Seventy-nine percent of Americans rate platforms' handling of harassment as only fair or poor.

5.4 CHILLING EFFECTS AND THE “POLITICS OF OUTRAGE”

Concerning the likelihood of decontextualised exposure, people may stop themselves from speaking freely or joking around. Institutions, too, can be affected. The police can act differently depending on expected reactions from the networked society.

5.5 LEGAL GROUNDING: REMEDIES, AND WHY THEY FAIL TO WORK

Although there are legal remedies available for victims, they will not be able to help much. The problem with these remedies is that they were made for a time when the media was much slower. An option here is defamation, although due to the ruling in *New York Times Co. v. Sullivan*,¹⁹ where the plaintiff has to prove that the defendant acted knowing that his statement was false or with reckless disregard for a statement's truthfulness, it becomes more difficult for plaintiffs who are considered public figures to recover damages, as it was their popularity that played a role in the distribution of

¹⁸Pew Research Centre, “The State of Online Harassment” (2021). Nationally representative survey of 10,093 U.S. adults, Sept. 8-13, 2020.

¹⁹ *New York Times Co. v. Sullivan*, 376 U.S. 254 (1964)

harmful information. In India, the IT Act Section 66E²⁰ implies that image capturing can be prosecuted, while similar torts can be found in other countries. However, generally, the accused will be legally responsible, not a multitude of users forming a coordinate harassment.

6. COUNTER-CURRENTS: TOWARD REPAIR AND REFORM

Sociologists may wrongly focus on the social phenomenon itself rather than the demands for which it exists: digital denunciation has exposed things such as harassment and discrimination that formal institutions have failed to address, thus helping digital vigilante justice seem a valid reaction to institutional failures. Social media trials can serve the purpose of causing differential and unjustifiable harm disconnected from the actual circumstances of the event.

Public opinion reflects this contradiction: according to a 2020 Pew Research Centre survey, 58% of American adults believe that online shaming is primarily about accountability, while 38% see it mainly as punishment of the innocent.

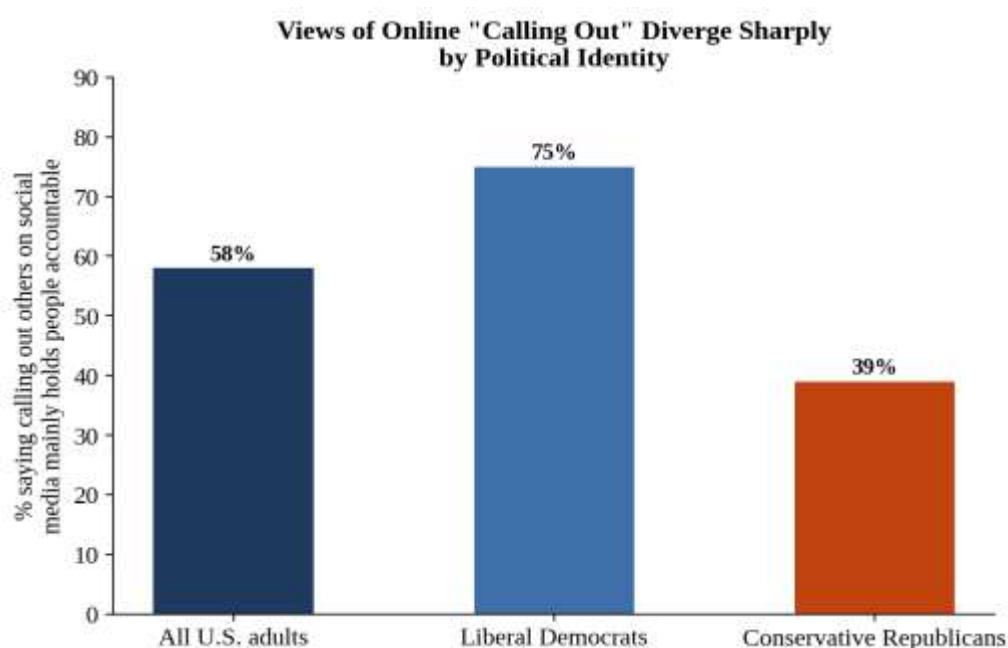


Figure 4: Views of online public shaming diverge sharply by political identity²¹

²⁰ IT Act Section 66E

²¹Pew Research Centre, "Americans and 'Cancel Culture': Where Some See Calls for Accountability, Others See Censorship, Punishment" (2021). Remaining respondents said calling out mainly punishes people who don't deserve it, or gave no answer.

Awareness has grown quickly: those reporting having heard “a fair amount” about “cancel culture” rose from 44% (2020) to 61% (2022).

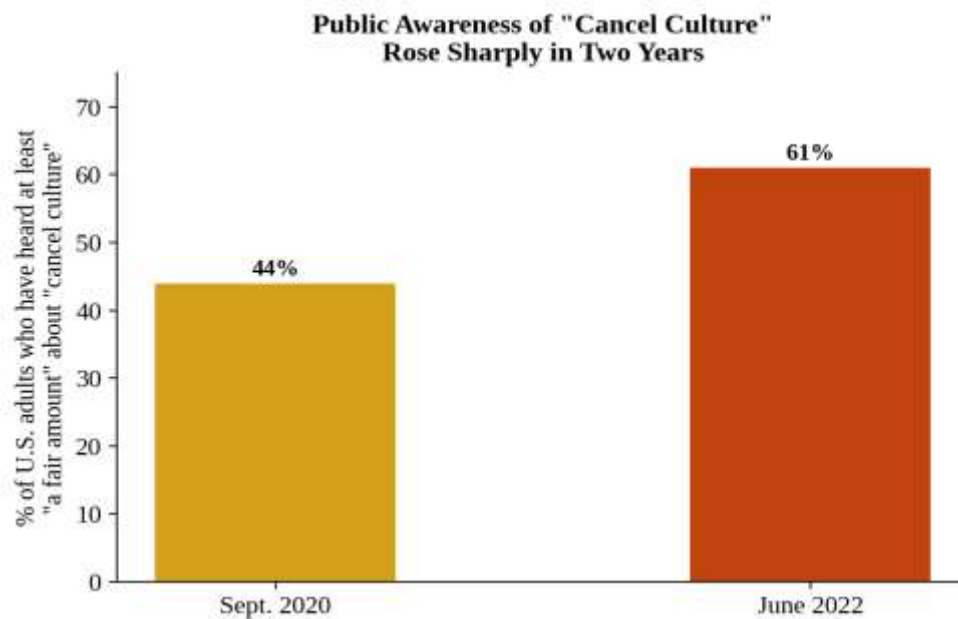


Figure 5: Public awareness of "cancel culture" rose sharply in two years²²

According to Kane's philosophical analysis, aligning with restoration principles must be pursued because shame-based call-outs muddle moral discourse and reinforce dichotomy. Kane alleges that accountability must aim for restoration instead of punishment. The propositions suggest a demagnifying algorithm for unproven accusations and the introduction of the concept of “the right to be forgotten” and demarcation between whistleblowing and organised harassment. The subsequent paragraphs discuss some of the measures already being implemented successfully as regulatory mechanisms.

7. REGULATORY AND INSTITUTIONAL RESPONSES: CENSOR BOARDS, CONTENT REGULATORS, AND PLATFORM SELF-GOVERNANCE

While Sections 2 to 6 have established that social media trials happen in a realm that is largely untouched by conventional law, this section discusses institutions that have recently emerged to fill this void. None has been created with the specific aim to conduct social

²²Pew Research Centre, “A Growing Share of Americans Are Familiar With ‘Cancel Culture’” (2022).

media trials, but all of them relate to the more general issue of harmful online content, of which shaming campaigns are only one type.

7.1 STATE REGULATORY AND STATUTORY GRIEVANCE BODIES

India's legal order is one of the highly sophisticated; therefore, it is noteworthy that the IT Rules 2021 require the “social media intermediaries” that meet a certain user-level threshold to appoint a grievance officer from the resident population of the country to deal with user complaints within the time period prescribed in the regulations.

The amendments introduced to the IT Rules in 2022 have established the Grievance Appellate Committees under Rule 3A by creating three-member committees that operate online and act as a quasi-tribunal with powers to consider appeals concerning decisions of the grievance officer and issue its ruling within a time frame that can be as short as 30 days.

Theoretically, the establishment of the Grievance Appellate Committees is beneficial since it provides affected individuals with the opportunity to get governmental support. However, practically, it may contradict the ruling of the Supreme Court in *Union of India v. R. Gandhi, President, Madras Bar Association*²³, which ruled that for bodies exercising judicial powers, it is necessary to be independent like the courts they intend to replace.

The Digital Services Act by the EU requires that each of its member countries to appoint a Digital Services Coordinator, hence giving the European Commission direct power over Very Large Online Platforms, and users are finally provided with the right to an internal complaints process and certified alternative dispute resolution.²⁴ The UK's Online Safety Act 2023 identifies Ofcom as the acting authority. The U.S. lacks a similar mandatory oversight organisation: Section 230²⁵ and the First Amendment limit judicial regulation to the relatively much narrower FTC consumer protection measures.

7.2 PLATFORM SELF-GOVERNANCE: META'S OVERSIGHT BOARD

Platforms have also created their own makeshift judicial entities. The Oversight Board of Meta, founded as a separate process that is funded by a Delaware trust with 20 international professionals in law and human rights²⁶, evaluates a limited number of disputed moderation

²³ Union of India v. R. Gandhi, President, Madras Bar Association, (2010) 11 SCC 1

²⁴ Regulation (EU) 2022/2065 (Digital Services Act).

²⁵ 47 U.S.C. § 230 (2018).

²⁶ U.N. Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights*, U.N. Doc. HR/PUB/11/04 (2011).

decisions in a year, taking into account the Community Standards of Meta and applicable international human rights laws, while operating under the established precedents of common law²⁷. Initial assessments have shown that it has reversed the majority of disputed decisions, showing that it does serve as some level of control over what a platform can do. However, the shortcomings are of a structural nature and include the fact that it processes only the tiniest fraction of all appeals it receives each year, that its jurisdiction extends only to this one company and that it reviews cases over a period of weeks and months; not really the right time to intervene in the crucial window of time when shaming campaigns inflict their greatest damage.

7.3 THE CENSORSHIP DILEMMA

All mechanisms discussed have a similar complication: an institution authorised to stop offensive content fast enough to prevent a social media pile-on is also capable of stopping legitimate speech with equal speed. India's experience is informative: Section 66A²⁸, which was recently struck down, criminalised “grossly offensive” speech. Indeed, as the Indian Supreme Court held in *Shreya Singhal*, it was leading to the censorship of “simple talk or advocacy” because vague notions of offence serve both the interests of the state and individuals. Therefore, the same can be said about any fast and low-profile content authority: the speed of operation makes its use against harassment possible while making the use of an agency against critical speech equally possible. None of the legal environments analysed so far has managed to resolve the dilemma, and it suffices to say that a co-regulatory framework still developing has emerged.

8. ANALYSIS

8.1 STRENGTHS OF THE PAPER

The paper stands out with its central choice of methodology, that is the combination of sociological mechanisms around relevant legal doctrine and its application in almost every section. Several works discussing shame on the Internet and cancel culture operate within a single format, such as the journalism of Ronson's *So You've Been Publicly Shamed* that illustrates the tragedy of the phenomenon through narratives but does not include comparative legal studies, while legal discussions on liability of platforms (debate on

²⁷ Evelyn Douek, *Facebook's "Oversight Board": Move Fast With Stable Infrastructure?*, 21 Yale J.L. & Tech. 1 (2020).

²⁸ IT Act Section 66A

Section 230 in the USA²⁹ or conversation on intermediary guidelines in India) talk about forces at play as a fact and not a subject of inquiry. But, by doing both things that is, applying Trottier's digital-vigilantism approach and discussing the *Shreya Singhal v. Union of India* case, using Brady et al.'s findings about virality and talking about the Digital Services Act, the paper brings a conclusion that social media trial is not only under-regulated in terms of the level of regulation used but is structurally resistant to the particular kind of regulation that existed in history and was applicable for media companies or cases observed in courts.³⁰

The second strength of the article is that it doesn't try to resolve the tension at its core. The authors never conclude that digital vigilantism is either a democratic tool correcting unresponsive institutions or simply a mob or act of violence. Instead, the authors leave both options open at once and argue that whether or not this is the case actually depends on situational conditions (who the target is, who first brings attention to the case, etc.) that are not really related to the original details of the crime. This is a more honest intellectual position than both extreme views, and the authors support it with empirical evidence, not just assertions. The Pew Research Centre's statistics about the fierce division in public opinion about whether or not calling out someone amounts to holding them accountable (75% of liberal Dems think it is the case versus 39% of conservative Republicans) prove that.³¹

Thirdly, it is noted that the treatment of regulatory responses in Section 7 is particularly forthright. Instead of presenting India's Grievance Appellate Committee, the EU's Digital Services Act, or Meta's Oversight Board as some solutions, the paper points out an inherent trade-off. Any institution that is fast enough to stop oncoming mischief has the same ability that allows it to suppress legitimate expression. The report anchors this dilemma by citing *Shreya Singhal's* ruling against the vagueness of India's Section 66A. This dilemma is perceived as open rather than closed, which is a more acceptable conclusion as compared to the usual "better regulation will fix it" view of papers of this kind.

8.2 LIMITATIONS OF THE PAPER

²⁹ Jack M. Balkin, *Free Speech in the Algorithmic Society*, 51 U.C. Davis L. Rev. 1149 (2018).

³⁰ William J. Brady, John T. Jost, Joshua A. Tucker & Jay J. Van Bavel, *Emotion Shapes the Diffusion of Moralised Content in Social Networks*, Proc. Nat'l Acad. Sci. U.S.A., 114(28), 7313-7318 (2017).

³¹ Ibid.

The empirical backdrop, as well as the comparative legal scope of the study, don't entirely corroborate each other. All five figures, along with any statistic that deals with harassment prevalence, severity and public perception, come from a single source, i.e the Pew Research Centre, which studies American adults only. However, the legal analysis does explicitly include comparative aspects as it highlights Indian, American and European doctrines. Unknown remains whether the underlying sociological patterns (growing harassment severity, the divide regarding “accountability” vs. “punishment”) also apply outside the United States, which makes it necessary to treat the empirical claims as much more reliable in relation to the U.S. only, while the comparative legal argument relies on a narrower and independently sourced evidential basis.

The second limitation is about the applicability of Trottier’s offence-taking, mediated-policing, and mediated-denunciation model. As the instances used to illustrate that model all belong to cases that are after an actual campaign is carried out, the paper cannot say why the majority of offensive or inappropriate posts do not result in digital vigilantism. The model explains past instances but does not predict which events will happen. Therefore, the case selection made by the paper does not let it move beyond that.

Third, the mechanisms that were studied in Section 4 put emphasis on the psychology of individuals and crowds. This is represented by the algorithms' amplification, context collapse, weaponised visibility, collective effervescence, and disinhibition. However, the incentives of the platform business were mentioned only in the legal sections about intermediary liability. A more detailed study would view advertising revenue as a mechanism that has the same sociological significance as context collapse or disinhibition, instead of including it in the legal area only.

In conclusion, this study's findings on effects focus on injury to individual victims. However, it does not consider another conceptually similar category: organisations such as universities, companies and public authorities that are subjected to organised online pressure and may be induced to change their policies as a result, which would require a different normative framework than the one presented for individual damage.

9. CONCLUSION

The concept of trial by social media is unique in sociology. It is a form of unofficial punishment based on its own concepts of virality and moral motivation. It draws from ancient traditions of boundary maintenance but its current form is given shape by contemporary aspects such as visibility as a weapon, collapse of context, algorithmic boost and digital permanence. Its results are not unequivocally positive or negative: it has helped reveal actual wrongdoings which formal institutions could not address, while resulting in excessive, discriminatory and often irreparable damage to people who had committed an act (real or imagined) that would not have warranted a comparable punishment under the existing system of formal legal principles.

Analysis in this paper demonstrates that existing doctrines were created for an old media environment characterised by the presence of known publications, registered proceedings, and mechanisms functioning on litigation terms. However, the technical terminology that characterises modern reality helps avoid the implementation of the described concepts referring to anonymous, decentralised and almost instantaneous damage. The regulatory bodies that are likely to be created nowadays, such as GAC in India, DSA architecture in Europe, or different platform control boards, seem to be authentic initiatives aimed at closing the gap mentioned above, although they share the same problems of how to balance the speed of completing tasks and the risks of becoming instruments of censorship. The issue of someone actively working on bridging the gap will remain relevant for both sociologists and legal scholars since, even in the conditions of modern platform design, one cannot neglect human behavioural tendencies, political economy issues related to platform design, or the need for accountability standards.

REFERENCES

I. Journal Articles

1. Ashworth, A., *Four Threats to the Presumption of Innocence*, 26 Int'l J. Evidence & Proof 215 (2022).
2. Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J., *Emotion Shapes the Diffusion of Moralised Content in Social Networks*, 114 Proc. Nat'l Acad. Sci. 7313 (2017).
3. Greer, C., & McLaughlin, E., *Trial by Media: Policing, the 24-7 News Mediasphere and the "Politics of Outrage,"* 26 Theoretical Criminology 84 (2022).
4. Kane, L. W., *Accountability and Community on the Internet: A Plea for Restorative Justice*, 37 J. Applied Phil. 594 (2020).
5. Marwick, A. E., & boyd, d., *I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience*, 13 New Media & Soc'y 114 (2011).
6. Suler, J., *The Online Disinhibition Effect*, 7 CyberPsychology & Behav. 321 (2004).
7. Trottier, D., *Digital Vigilantism as Weaponisation of Visibility*, 30 Phil. & Tech. 55 (2017).
8. Trottier, D., *Denunciation and Doxing: Towards a Conceptual Model of Digital Vigilantism*, 21 Global Crime 196 (2020).
9. Various Authors, *Introducing Digital Vigilantism*, 21 Global Crime 189 (2020).
10. Wahyudi, D., Sujoko, A., & Ayub, Z. A., *The Presumption of Innocence: Interpretation and Application in Online Journalism*, 52 Informasi 215 (2022).
11. Evelyn Douek, *Facebook's "Oversight Board": Move Fast With Stable Infrastructure?* 21 Yale J.L. & Tech. 1 (2020).
12. Jack M. Balkin, *Free Speech in the Algorithmic Society*, 51 U.C. Davis L. Rev. 1149 (2018).

II. Books

1. Citron, D. K., *Hate Crimes in Cyberspace* (Harv. U. Press 2014).
2. Ronson, J., *So You've Been Publicly Shamed* (Riverhead Books 2015).

III. Reports / Web Sources

1. Pew Rsch. Ctr., *The State of Online Harassment* (Jan. 13, 2021), <https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/>.
2. Pew Rsch. Ctr., *Americans and "Cancel Culture": Where Some See Calls for Accountability, Others See Censorship, Punishment* (May 19, 2021),

<https://www.pewresearch.org/internet/2021/05/19/americans-and-cancel-culture-where-some-see-calls-for-accountability-others-see-censorship-punishment/>.

3. Pew Rsch. Ctr., *A Growing Share of Americans Are Familiar with “Cancel Culture”* (June 9, 2022), <https://www.pewresearch.org/short-reads/2022/06/09/a-growing-share-of-americans-are-familiar-with-cancel-culture/>.

IV. Cases

1. Nebraska Press Ass’n v. Stuart, 427 U.S. 539 (1976).
2. New York Times Co. v. Sullivan, 376 U.S. 254 (1964).
3. Sahara India Real Estate Corp. Ltd. v. SEBI, (2012) 10 SCC 603.
4. Shreya Singhal v. Union of India, AIR 2015 SC 1523.
5. Union of India v. R. Gandhi, President, Madras Bar Ass’n, (2010) 11 SCC 1.

V. Legislation & Regulations

1. Communications Decency Act of 1996, 47 U.S.C. § 230 (U.S.).
2. Contempt of Court Act 1981 (U.K.).
3. Contempt of Courts Act, 1971 (India).
4. Information Technology Act, 2000, §§ 66E, 79 (India).
5. Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, as amended 2022 (India).
6. Online Safety Act 2023 (U.K.).
7. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services (Digital Services Act).