

Is social media a teaching tool for fostering racial tolerance in western society?

Huiyu Zhou

University College London (UCL), WC1H, 0AL, London, UK,

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This study questions whether social media functions as a teaching tool for fostering racial tolerance in Western society. Using Critical Race Theory (CRT) and Social Media Influence Theory, the research examines how social hierarchy, algorithmic designs, platform governance, normalisation, psychological elements, and surface effects shape racial discourse. However, the counterarguments are that rules in social media encourage fairness and communication among races as it gives voice to victims. The argumentation is that racial prejudice in Western online environment is even fostered and exaggerated by social media intentionally, instead of reflecting racial biases. It will thus recommend solutions related with technological and governmental remedies.

Keywords: social media, racial tolerance, critical race theory, social media influence theory, western society

Online publication: August 26, 2025

1. Introduction

Racial discourse has been altered by social media, which simultaneously provides a forum for under-represented views and reinforces structural racial inequality^[1]. Because of this dual nature, important concerns are brought up, such as how much digital platforms support racial justice and how their fundamental structures influence racial narratives. This essay argues that racial prejudice in Western cultures is intentionally fostered and exaggerated by social media, rather than just reflecting racial biases. It will begin with the literature review section summarising the literature on the Critical Race Theory, Social Media Influence Theory, current research on race in social media, and research gaps. The argument part will investigate how digital platforms promote racial enmity by examining social hierarchy, algorithmic designs, platform governance, normalisation, psychological elements, and surface effects. It will then list counterarguments that social media encourages racial fairness and inter-race communication through setting regulations and giving voice to victims while critically evaluating their shortcomings. The conclusion will examine possible technological and governmental remedies.

2. Literature Review

The literature on the Critical Race Theory, Social Media Influence Theory, current race-related social media research and research gaps is summarised in this section. A conceptual framework for comprehending racial dynamics in social networking sites is offered by Critical Race Theory (CRT). CRT holds that xenophobia is a systematic, established aspect

of Western cultures that is ingrained in organisations and structures rather than a collection of discrete episodes^[2,3]. Because racism is systematic, it employs coordinated, covert acts to uphold racial disparities even when there are no overt signs of racism^[4,5]. CRT emphasises that racial advancement usually coincides with dominant white groups' goals, a process called interest convergence (ibid.). Therefore, advancements in racial equality are viewed as both morally right and as pragmatic compromises that benefit the rich. Furthermore, CRT emphasises that race is a socially created concept impacted by historical, political, and social variables rather than a fixed biological reality.

Another indispensable facet CRT is the counternarratives of the marginalized groups themselves, which can be used to dispute the predominating racial narratives. Zillmann's (2002)^[6] exemplification theory states that media frequently employs representative examples that arouse viewers' intense emotions, hence affecting their attitudes and perceptions. To combat false information and divisive narratives that are common in the digital sphere, Social Media Influence Theory emphasises the significance of digital literacy, open algorithms, and ethical content monitoring^[7]. There are still several study gaps even though the existing work offers insights into how social media influences racial discourse. First, the part that platform governance and regulatory frameworks play in racist debate on social media is not well understood. Second, not enough research has been done on how social media platforms' built-in financial incentives and business needs affect racial debate. Finally, even though counter speech has become a viable tactic for addressing online racism, most current research is on short-term interventions, with little understanding of the long-term effects of these initiatives. The argumentation part will fill in these gaps by examining the effects of financial incentives, governance frameworks, and ongoing anti-racist initiatives.

3. Arguments

3.1. Political and Social Factors

Regarding the context of the western countries, distinguished by a democratic political climate that protects fundamental liberties and rights^[8]. As a result, freedom is frequently valued more highly than control in the political economy of the internet, frequently at the price of regulatory monitoring^[9]. For democratic countries and the capitalist economies that support them, this absence of regulation poses serious difficulties as they attempt to strike a balance between the necessity to prevent harmful material and the ideals of free expression^[10]. Critical Race Theory (CRT) also claims how institutional norms, not just personal prejudices, are what sustain structural racism. Silva (2018) emphasises how systematic racism still shapes social hierarchies and possibilities in modern society by defining white supremacy as a racialised society made up of all the social interactions and customs that uphold white privilege. In this background, the propensity to place blame on marginalised groups by blaming the victim is one of the main elements of the contemporary, indirect type of racial discrimination, which is sometimes referred to as "colourblind racism"^[11].

3.2. Algorithmic Bias

By selecting information that supports users' pre-existing opinions, these algorithms, which are intended to tailor user experiences, reinforce ideological divides and restrict exposure to opposing views^[12]. Social media platforms hinder true racial understanding by establishing knowledge bubbles that separate users from competing viewpoints rather than encouraging open discourse^[13]. Dominant narratives thus gain more attention, while marginalised viewpoints are often silenced. Algorithms prioritise conflict-driven narratives above productive dialogues, which not only reinforces bias but also speeds up the dissemination of racially offensive information. As social media sites use engagement numbers, anything that incites anger and division is more likely to be shared. As antagonistic viewpoints gain hold and voices calling for racial understanding fail to be heard, this tendency distorts public opinion^[14]. Racial animosity is further exacerbated by social media algorithms that put interaction ahead of morality^[15].

3.3. Platform and Designers

The structural layout of platforms and inadequate regulation also play a part in the spread of hatred. The arbitrary attitude taken by the International Telecommunications Union (ITU), which oversees the global communications infrastructure, regarding the messages that are sent via its system has drawn criticism^[16]. For example, Instagram and Snapchat provide filters that soften the skin tones of people who are not white, therefore perpetuating Eurocentric ideals of beauty^[17]. According to Angwin and Parris (2016)^[18], Facebook has also come under fire for enabling marketers to exclude people from ad campaigns since ethnic affinities. Moreover, the nature of social media promotes anonymity, disengagement, and disinhibition, enabling users to voice extreme viewpoints without worrying about the repercussions in the actual world^[19]. Another problem, as noted by Chaudhry and Gruz (2020)^[20], is the over-reliance on individual users to combat racism. This assumption, however, disregards CRT's claim of interest convergence because many users do not have the personal will to delve extensively into topics that do not immediately impact their interests or social identity.

3.4. Normalization of Racism

Furthermore, when hate speech and racially disparaging information are often shared on social media, they gradually gain acceptance^[21]. The dissemination of harmful jokes and racialised humour is one of the most common ways that racism is normalised online. Prejudice may be disguised as humour or free speech through subtle strategies like the weaponization of memes and the employment of fake identities^[22,23]. A culture that trivialises racist insults by characterising them as just jokes and portraying people who take offence as overly sensitive or against free speech is fostered by this digital backstage environment^[24]. Under the guise of innocuous amusement, social media's constant dissemination of racist humour desensitises users to racial prejudice and reinforces systemic prejudices^[25].

3.5. The Superficial Impact of Online Activism

Genuine racial understanding is thwarted by the fact that many platforms place more emphasis on surface-level interaction likes, shares, and trending hashtags than on real conversation^[26]. Furthermore, in terms of shaping racial attitudes, in-person interactions remain more significant than online engagement. Tao and Fisher's (2022)^[27] idea of the negative feedback loop makes clear the limitations of internet activism. Jakubowicz (2017)^[28] supports this assertion by pointing out that many anti-racism activities fall short in their efforts to overcome structural biases. This criticism is expanded upon by Coles and Lane (2023)^[29], who contend that a lot of interventions fall short because they depend on one-time training sessions or generic messages instead of ongoing interaction with impacted populations.

3.6. Psychological Barriers to Racial Progress

The psychological aspects of online interactions serve to further strengthen racism's existence on social media. Research indicates that exposure to discriminatory information frequently results in emotional discomfort, which lowers people's capacity for sympathetic participation in racial discourse^[30]. This phenomenon is explained by Critical Race Theory (CRT), which holds that confronting racism inevitably exposes deeply ingrained systemic beliefs that favour dominant groups, making it psychologically taxing for people to critically engage with racial injustices. Furthermore, according to Social Media Influence Theory, algorithmic curation amplifies emotional and sensationalist content, including discriminatory material, resulting in heightened emotional responses that wear users out and decrease their willingness to engage in productive discourse.

4. Counterarguments

4.1. Establish Norms

The opposing opinion that social media is essential for creating standards for racial discourse since it encourages marginalised perspectives and deters discriminatory behaviour. This is exemplified by movements such as

#BlackLivesMatter, which have significantly influenced public conversations on systemic racism^[31]. However, norm-setting's efficacy differs depending on the platform. Due to uneven execution of moderation measures, racist language is occasionally not just accepted but actively propagated^[32]. While some platforms have robust anti-hate speech policies, others do not take swift action to delete discriminatory information, allowing it to spread before any action is done^[33]. These organisations frequently create high-engagement material intended to elicit responses and boost exposure to influence platform algorithms^[34]. Additionally, the creation of echo chambers, where users primarily engage with groups that have similar opinions, limits the spread of opposing perspectives and reduces the overall impact of efforts to create norms.

4.2. Amplifying the Voices of Victims

Another rebuttal might be that social media gives victims of racial injustice a vital forum to talk about their own stories, which promotes awareness, empathy, and unity^[35]. Videos, pictures, and first-person accounts assist viewers empathise with victims' experiences by personalising the hardships faced by ethnic minorities. Visibility by itself, meanwhile, does not ensure transformation. Re-traumatizing victims by recurrent reliving of their traumas is a danger associated with continuous exposure to upsetting information. Once their tales become viral, victims frequently have little influence over how they are exploited, which can result in emotional fatigue and the exploitation of their suffering^[36]. Furthermore, audiences may become desensitised because of frequent exposure to upsetting material, which gradually lessens the emotional effect of racial violence^[37]. Excessive exposure might normalise racial violence as an inevitable reality rather than an issue that requires immediate attention, which can elicit indignation and action^[38].

5. Conclusion

In summary, social media does more to encourage racial animosity than to develop racial concord. The propagation of false information, the normalisation of racist speech, and the reinforcing of prejudices are all facilitated by its algorithmic frameworks, financial incentives, and governance shortcomings. Social media will continue to be a major source of racial prejudice rather than a force for good in Western culture unless there are structural reforms made to platform regulation, content control, and digital literacy. Governments and advertisers should put financial disincentives on platforms that benefit from racist participation, either through penalties, taxes, or the removal of advertising income from non-compliant corporations, as financial incentives frequently fuel the persistence of bad material. Together, these actions can help turn social media into a platform that encourages racial tolerance rather than escalating hate.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Pantti M, Nelimarkka M, Nikunen K, et al., 2019, The meanings of racism: Public discourses about racism in Finnish news media and online discussion forums. *European Journal of Communication*, 34(5): 503-519.
- [2] Delgado R, Stefancic J, 2023, *Critical race theory: An introduction*. NyU press.
- [3] Taylor E, 1998, A primer on critical race theory: Who are the critical race theorists and what are they saying? *The journal of blacks in higher education*, (19): 122-124.
- [4] Ladson-Billings G, 2021, Critical race theory—What it is not! In *Handbook of critical race theory in education*. Routledge. 32-43.

- [5] Delgado R, Stefancic J E, 2000, *Critical race theory: The cutting edge*. Temple University Press.
- [6] Zillmann D, 2002, Exemplification theory of media influence. In *Media effects*. Routledge. 29-52.
- [7] Rachmad Y E, 2023, Social Media Influence Theory. *Technology and Society Perspectives (TSP)*. 1(2): 56–65.
- [8] Weber A, 2009, *Manual on Hate Speech*. Council of Europe Publishing.
- [9] Klein A, 2017, *Fanaticism, racism, and rage online: Corrupting the digital sphere*. Springer.
- [10] McGonagle T, 2012, Minorities and Online ‘Hate Speech’: A Parsing of Selected Complexities. *European Yearbook of Minority Issues Online*, 9(1): 419-440.
- [11] Silva E B, 2018, *Racism without racists*. Rowman & Littlefield.
- [12] Rodilosso E, 2024, Filter bubbles and the unfeeling: How AI for social media can foster extremism and polarization. *Philosophy & Technology*, 37(2): 71.
- [13] Zimmer F, Scheibe K, Stock M, et al., 2019, Fake news in social media: Bad algorithms or biased users?. *Journal of Information Science Theory and Practice*, 7(2): 40-53.
- [14] Stray J, Iyer R, Puig L H, 2023, The algorithmic Management of Polarization and Violence on social media. *PNAS Nexus*, 2(6): 182.
- [15] Walsh J P, 2020, Social media and moral panics: Assessing the effects of technological change on societal reaction. *International Journal of Cultural Studies*, 23(6): 840-859.
- [16] International Telecommunications Union, 2017, What does ITU do? Retrieved January 21, 2017, from <http://www.itu.int/en/about/Pages/whatwedo.aspx>.
- [17] Jerkins M, 2015, The quiet racism of Instagram filters. *Racked*. <https://www.racked.com/2015/7/7/8906343/instagram-racism>.
- [18] Angwin J, Parris T J, 2016, Facebook lets advertisers exclude users by race. *ProPublica*. <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>.
- [19] Buckels E E, Trapnell P D, Paulhus D L, 2014, Trolls just want to have fun. *Personality and individual Differences*, 67: 97-102.
- [20] Chaudhry I, Gruzd A, 2020, Expressing and challenging racist discourse on Facebook: How social media weaken the “spiral of silence” theory. *Policy & Internet*, 12(1): 88-108.
- [21] Tsesis A, 2001, Hate in cyberspace: Regulating hate speech on the Internet. *San Diego L. Rev.*, 38, 817.
- [22] Lamerichs N, Nguyen D, Melguizo M C P, et al., 2018, Elite male bodies: The circulation of alt-Right memes and the framing of politicians on Social Media. *Participations*, 15(1): 180-206.
- [23] Farkas J, Schou J, Neumayer C, 2018, Cloaked Facebook pages: Exploring fake Islamist propaganda in social media. *New Media & Society*, 20(5): 1850-1867.
- [24] Pérez R, 2017, Racism without hatred? Racist humor and the myth of “colorblindness”. *Sociological Perspectives*, 60(5): 956-974.
- [25] Feagin J R, 2020, *The white racial frame: Centuries of racial framing and counter-framing*. Routledge.
- [26] Page R, 2012, The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags. *Discourse & communication*, 6(2): 181-201.
- [27] Tao X, Fisher C B, 2022, Exposure to social media racial discrimination and mental health among adolescents of color. *Journal of youth and adolescence*, 51(1): 30-44.
- [28] Jakubowicz A, 2017, Alt Right White Lite: trolling, hate speech and cyber racism on social media. *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, 9(3): 41-60.
- [29] Coles S M, Lane D S, 2023, Making the impossible possible? Framing confrontations of racism on social media as norm-setting. *New Media & Society*, 14614448231208707.
- [30] Sawyer P J, Major B, Casad B J, et al., 2012, Discrimination and the stress response: Psychological and physiological consequences of anticipating prejudice in interethnic interactions. *American journal of public health*, 102(5): 1020-1026.
- [31] Dunivin Z O, Yan H Y, Ince J, et al., 2022, Black Lives Matter protests shift public discourse. *Proceedings of the National*

Academy of Sciences, 119(10): e2117320119.

- [32] Kretzmer D, 1986, Freedom of speech and racism. *Cardozo L. Rev.*, 8, 445.
- [33] Gagliardone I, Gal D, Alves T, et al., 2015, Countering online hate speech. *Unesco Publishing*.
- [34] Bushe G R, Lewis S, 2023, Three change strategies in organization development: data-based, high engagement and generative. *Leadership & Organization Development Journal*, 44(2): 173-188.
- [35] Nikunen K, 2018, Media solidarities: Emotions, power and justice in the digital age.
- [36] Hoskins A, 2020, Media and compassion after digital war: Why digital media haven't transformed responses to human suffering in contemporary conflict. *International Review of the Red Cross*, 102(913): 117-143.
- [37] Scharrer E, 2008, Media exposure and sensitivity to violence in news reports: Evidence of desensitization? *Journalism & Mass Communication Quarterly*, 85(2): 291-310.
- [38] Carter C, 2003, *Violence and the Media*. McGraw-Hill Education (UK).

Publisher's note

Whioce Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.