Harvard Kennedy School Misinformation Review[1]
November 2026, Volume 7, Issue 7
Creative Commons Attribution 4.0 International (CC BY 4.0)
Reprints and permissions: misinforeview@hks.harvard.edu
DOI: https://doi.org/10.37016/mr-2020-215
Website: https://misinforeview.hks.harvard.edu

Research Article

Racially biased disinformation as structural violence: A discursive operations framework from Brazil

On March 14, 2018, Marielle Franco, a Black, queer city councilwoman from Rio de Janeiro, was assassinated. Within hours, false claims that she was "married to a drug trafficker" began circulating; formally debunked, they recirculated two years later. The case exemplifies a structural pattern: in Brazil, disinformation targeting Black individuals who die amid State violence consistently substitutes identities, fabricates criminal histories, and defames private lives to retroactively legitimize racial violence. Analyzing 3,674 fact checks from two IFCN-certified Brazilian agencies (2018–2023), this article proposes a replicable five-operation discursive framework for understanding how racially biased disinformation operates and compounds its harm.

Authors: Edson Mendes Nunes Junior (1), Luan Cardoso Ferreira (2)
Affiliations: (1) Department of Political Studies, Latin American Faculty of Social Sciences (FLACSO Ecuador), Ecuador, (2) Center for Social Theory and Latin America Studies (NETSAL), State University of Rio de Janeiro (UERJ), Brazil
How to cite: Nunes Junior, E. M., & Ferreira, L. C. (2026). Racially biased disinformation as structural violence: A discursive operations framework from Brazil. *Harvard Kennedy School (HKS) Misinformation Review*, *7*(7).
Received: June 1st, 2026. Accepted: October 16th, 2026. Published: November 2nd, 2026.

Research questions

- What discursive operations does racially biased disinformation deploy against Black individuals and groups in Brazil, and how do these operations combine?
- What are the structural consequences of compound operations for disinformation harm and the dynamics of refutation?
- How does the Brazilian case operationalize existing frameworks of racialized disinformation in a Global South context?
- What methodological distinction is necessary between racially biased disinformation and disinformation merely associated with racial cases?

Essay summary

- We analyzed 3,674 fact-checks published by *Agência Lupa* and *Aos Fatos* (2018–2023), identified 192 racially relevant items through a dictionary-based filter, and confirmed 34 as racially biased disinformation (RBD) through manual coding (Table 1).

[1] *A publication of the Shorenstein Center on Media, Politics and Public Policy at Harvard University, John F. Kennedy School of Government.*

- We propose a discursive operations framework comprising five analytically distinct but empirically combinable operations: identity substitution (SI), criminalization (CR), moral defamation (DM), relational association (AR), and decontextualization (DC), defined in full in Table 2.
- 72% of confirmed cases (24/34) involve two or more simultaneous operations. The dominant compound pattern is SI+CR (*n* = 10, 29.4%), followed by DC+CR (*n* = 8, 23.5%), suggesting compound operations are the rule, not the exception.
- We introduce the term *operational complexity*, the structural asymmetry that multiple simultaneous operations may produce, in which disinformation compounds in harm while refutation remains linear, a theoretical implication that future research should test empirically.
- We formalize a methodologically necessary distinction between racially biased disinformation (RBD), which deploys operations against Black individuals or groups as primary targets, and disinformation associated with racial cases (DARC), in which a racial event serves as political backdrop while the primary operation targets white actors.
- Marielle Franco is the most targeted individual (7/34 cases), with a diversified ecosystem deploying all five operations, including a novel sub-type of AR as victim-to-perpetrator inversion.
- Facebook was the dominant circulation platform (73.5%). Criminalization (CR) appeared in 76.5% of all confirmed cases, confirming its centrality to the punitivist legitimation of racial violence.

Implications

Racialized disinformation and the gap in its operationalization

The concept of racialized disinformation is established in the literature. The Media Manipulation Casebook (2021) defines it as content deliberately false or misleading and targeting wedge issues related to race, racial justice, or communities of color, grounding it in Omi and Winant's (1994) notion of racialization as a sociohistorical process through which racial categories are created, inhabited, and transformed. Kuo and Marwick (2021), publishing in this journal, argue that disinformation scholarship must engage with history, power, and racial formation, and that treating disinformation as an ideologically neutral "toxin" renders its origins and resonances unintelligible. Reddi, Kuo, and Kreiss (2021) developed the related concept of "identity propaganda," narratives that strategically exploit identity-based differences to maintain hierarchies of power. In the U.S. context, Freelon, McIlwain, and Clark (2016) documented how disinformation functioned as counter-mobilization against Black Lives Matter, demonstrating the operation of racial narratives within disinformation ecosystems. Gehrke and Amit-Danhi (2025), also in this journal, have recently extended this agenda to gendered disinformation, observing that only 7% of papers at leading communication conferences address gender or race in connection with disinformation.

What the existing literature does not yet provide is a systematic framework for understanding how racially biased disinformation compounds its harm and resists refutation, specifying why compound operations persist, despite correction, in ways that single-operation claims do not. This article addresses that gap. We do not introduce the concept of racialized disinformation; we operationalize it for the Global South through a framework that makes it empirically tractable, codeable across cases, and theoretically connected to the structures of racial inequality on which each operation depends.

Brazil as a theoretically strategic case

Brazil is not merely an empirical setting but an analytically productive case. With approximately 113 million Black and mixed-race inhabitants, representing 55.5% of its population (IBGE, 2023), Brazil has one

of the largest African-descendent populations in the world outside the African continent. The country's racial formation is defined by the ideology of racial democracy, a set of ideas that presents Brazil as a harmonious, racially mixed society and thereby naturalizes inequality, masks exclusion, and blocks Black political consciousness, historically suppressing recognition of structural racism while maintaining racialized hierarchies in practice (Almeida, 2018; Moura, 1994): 67% of the incarcerated population in Brazil is Black (INFOPEN, 2021), and Black Brazilians constitute the overwhelming majority of victims of lethal police violence, concentrated in favela and peripheral territories (Fórum Brasileiro de Segurança Pública, 2023). This structure intersects with what Wacquant (2007) terms *punitivism*, in which the state, as welfare protections retreat, increasingly manages the social insecurity of marginalized populations through policing and incarceration rather than through redistributive policies such as income support, housing, or employment. In Brazil, punitivism is mediatized: sensationalist police journalism constructs the "bandido" as young, Black, and peripheral (Romão, 2013), and the affective logics of Brazilian media, which address audiences through emotion and spectacle rather than argument, amplify its reach by privileging fear over reasoned deliberation (Sodré, 2006). When social media expanded through Brazil's unprecedented WhatsApp penetration, this punitivist logic migrated into digital disinformation. What Mbembe (2018) calls *necropolitics*, the sovereign power to decide who may live and die, finds in disinformation a new mechanism: after a killing, false narratives recast the victim as a criminal, destroying their memory and retroactively legitimizing lethal State action. The period 2018–2023 makes this dynamic unusually visible: it spans the assassination of Marielle Franco, a Black city councilwoman from a Rio de Janeiro favela (March 2018); the COVID-19 pandemic; the Jacarezinho police operation in Rio de Janeiro (28 killed, May 2021), the deadliest in the city's history; and the January 8, 2023, in which supporters of former president Jair Bolsonaro stormed Congress, the Presidential Palace, and the Supreme Court in Brasília after Lula's inauguration. Over the same years, legislative attention to disinformation intensified but framed it primarily as a threat to elections and political processes, leaving its racial targeting largely unaddressed (Anastácio, 2024). Across all these moments, the same five operations recurred—identity substitution (SI), criminalization (CR), moral defamation (DM), relational association (AR), and decontextualization (DC)—suggesting that racially biased disinformation in Brazil is not episodic but structural.

Operational complexity and the asymmetry of refutation

The central theoretical implication of this article is that racially biased disinformation may be compounding in harm while refutation remains linear, an asymmetry suggested by the operational architecture of compound claims, that is, narratives that combine two or more discursive operations, and one that warrants direct empirical testing. When a single discursive operation is deployed, refutation is tractable: prove the image is false, prove the criminal attribution is fabricated. When two or more operations are simultaneously deployed, occurring in 72% of our confirmed cases, each operation demands its own refutation, and debunking one leaves the other intact to keep the claim circulating—since correcting one element does not remove the content from circulation, and platforms do not amplify corrections at the rate at which original claims spread. The refutation burden therefore grows with every added operation, while the claim's reach does not diminish proportionally. We term this *operational complexity*.

The SI+CR compound, in which a false image (SI) and a criminal attribution (CR) reinforce each other, is paradigmatic. Proving a substituted image false leaves the criminal attribution still circulating, and vice versa. Each element reaches a different audience segment through a different cognitive channel: the image through visual immediacy, the criminal attribution through punitivist common sense. Both can be formally refuted and still persist, because each draws on a different pre-existing racial infrastructure. Ecker and colleagues (2022) document the psychological persistence of misinformation beliefs after

correction; operational complexity explains why this persistence is structurally amplified in compound racial disinformation, where the architecture of the claim is designed, whether intentionally or not, to outlast any single act of debunking.

Evidence

The four findings presented below draw on five discursive operations identified through our coding process and defined in full in Table 2 (Methods): *identity substitution* (SI), the use of a false image or identity to misrepresent a Black victim as a criminal actor; *criminalization* (CR), fabricated or decontextualized attribution of criminal history or behavior; *moral defamation* (DM), attacks on the victim's private life or relationships; *relational association* (AR), delegitimization through fabricated associations with criminal actors or racially coded territories; and *decontextualization* (DC), factually accurate content placed in a misleading frame to criminalize Black communities. Operations are not mutually exclusive; compound cases are coded with "+" (co-primary) or ">" (primary > secondary).

Finding 1: SI+CR is the dominant compound operation, extending a previously documented pattern of post-mortem criminalization.

The most frequent compound pattern, SI+CR—*identity substitution* (SI), the use of a false image or identity to misrepresent a Black victim as a criminal actor, combined with *criminalization* (CR), the fabricated or decontextualized attribution of criminal history or behavior (n = 10, 29.4% of the 34 confirmed cases)—extends a pattern first documented in Brazilian police violence cases in 2014–2015 (Nunes Junior, 2021). In the Jacarezinho case (May 2021), photographs of convicted prisoners from Rio Grande do Norte, a state in northeastern Brazil, were circulated on social media with captions presenting them as the 28 residents killed in the police operation in Rio de Janeiro. The groups were unrelated, but the false captions recast the dead as convicted criminals, which functioned as a retroactive justification for the killings. The mutual reinforcement between operations is structurally significant: the substituted images make the criminal attribution visually plausible, and the criminal attribution provides the interpretive frame through which audiences accept the false images. Neither element alone would be as resilient. Together, they create a cluster demanding two different categories of evidence to dismantle, and the window of reputational damage passes before both are addressed. Across all SI+CR cases, the subjects of false images were Black and from peripheral territories, because the visual substitution depends on the low public recognition of peripheral individuals and on the punitivist common sense that makes criminal associations with Black bodies feel instinctively believable and accepted without conscious scrutiny (Figure 1).

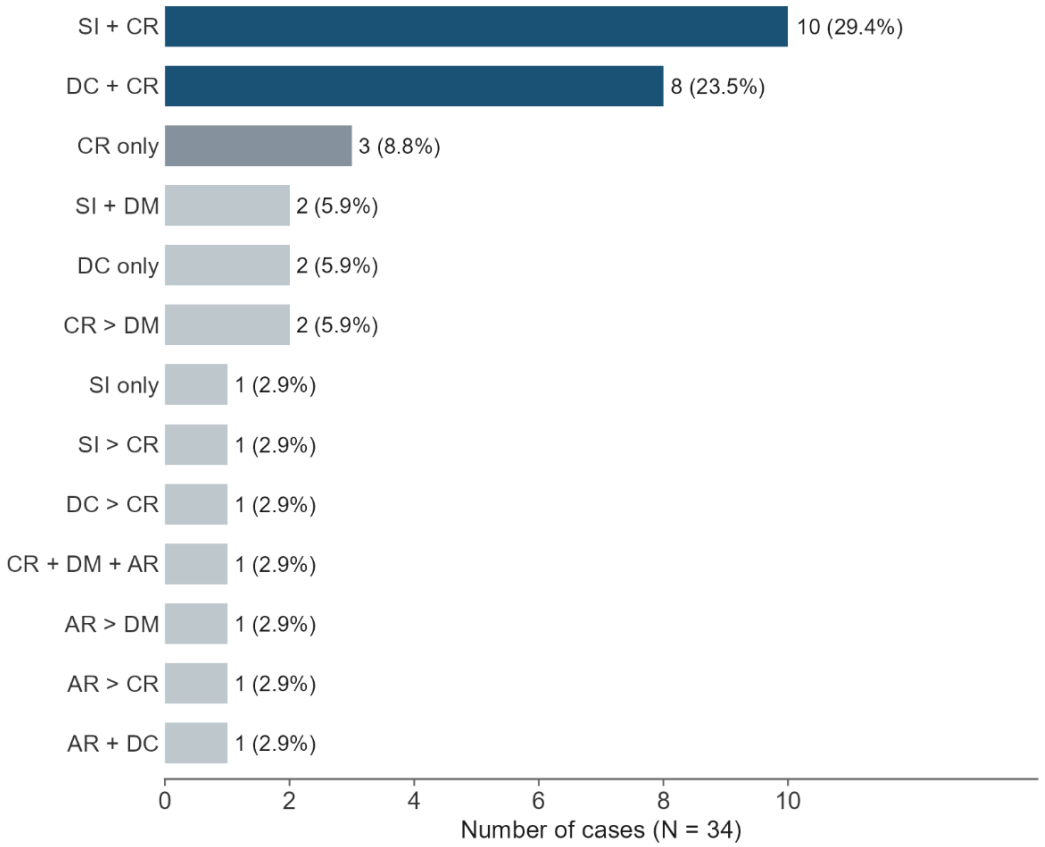


Figure 1. Distribution of discursive operations in confirmed cases of racially biased disinformation (*N *= 34). *Operations: SI = identity substitution; CR = criminalization; DM = moral defamation; AR = relational association; DC = decontextualization. Darker bars indicate dominant patterns. Operations are not mutually exclusive; compound cases are coded with "+" (co-primary) or ">" (primary > secondary). Source: authors' manual coding of Agência Lupa and Aos Fatos fact-checks, 2018–2023.*

Finding 2: DC+CR constitutes a second, analytically distinct cluster deploying true content to criminalize Black movements and delegitimize judicial protection.

The second most frequent pattern, DC+CR (*n* = 8, 23.5%), does not appear in prior systematic analyses of racially biased disinformation in Brazil (Nunes Junior, 2021) and represents this article's most novel empirical contribution. In these cases, factually accurate content is removed from its original context and reframed to support the narrative that Black communities are inherently dangerous or that protections extended to them produce disorder. Four cases share an identical structure: old *baile funk* videos or clips from fictional television series were circulated as evidence of disorder following the 2021 Supreme Court ruling that suspended police operations in Rio's favelas during the COVID-19 pandemic. Because the underlying content is not fabricated, DC+CR is technically harder to refute as "false" and tends to persist longer in circulation. An old looting photograph shared as evidence that Black Lives Matter protests produced violence in Brazil (Aos Fatos, 2020) extends this logic transnationally: DC+CR functions not only against domestic Black communities but as a counter-mobilization tool against transnational Black political movements, consistent with how disinformation operates in U.S. racial justice contexts (Freelon et al., 2016).

Finding 3: The Marielle Franco disinformation ecosystem reveals differentiated operational architecture and the persistence of post-mortem racial delegitimization.

Marielle Franco is the most targeted individual in the corpus (*n* = 7, 20.6%), and the confirmed cases against her are not operationally repetitive but differentiated across five configurations, each targeting a distinct dimension of her identity. AR+DM cases, combining *relational association* (AR) and *moral defamation* (DM), constructed her as morally undeserving of political grief; unlike CR, which positions a victim as a threat requiring violent response, DM positions her as personally compromised and unworthy of collective solidarity. The most extreme case, a false claim that her aide's fingerprints were found in the assassins' vehicle, deployed what we term AR as victim-to-perpetrator inversion: the disinformation falsely implicated her own support network in her murder, representing the most severe form of AR documented in this corpus. The "Marcinho VP" relational association claim (a false allegation linking Franco to a Rio drug trafficker) was fact-checked twice, in 2018 and again in 2020, confirming Pollak's (1989) framework of memory as a contested field: racially biased disinformation functions not as a discrete event but as sustained memory contestation, reconstituting each time its target re-enters public attention (Figure 2).

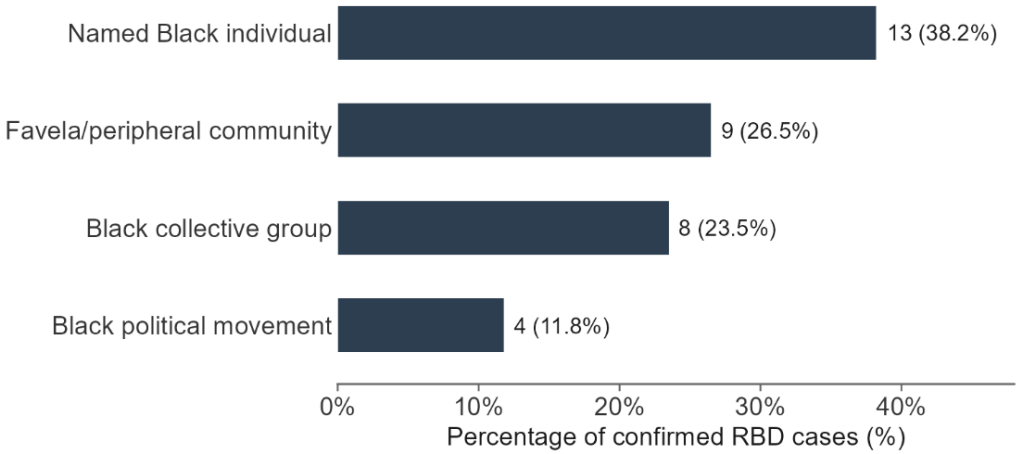


Figure 2. Target categories in confirmed cases of racially biased disinformation (N = 34). *RBD = racially biased disinformation. Source: authors' manual coding, 2018–2023.*

Finding 4: The distinction between racially biased disinformation (RBD) and disinformation associated with racial cases (DARC) reveals a parallel ecosystem.

A methodologically necessary finding concerns what is not in the corpus. Of the 155 excluded items, a minimum of 23 (and likely substantially more) involved false claims about white political actors using the Marielle Franco investigation as context. For example, one excluded item falsely claimed that a white federal deputy had leaked confidential details of the Marielle Franco murder investigation to protect an ally; the racial event, Franco's assassination, supplied the political backdrop, but the disinformation's target and the reputational harm it sought to inflict were directed at the deputy, a white actor, using the *identity substitution* (SI) operation to substitute a fabricated leaked document for a real one. This is the defining feature that separates disinformation associated with racial cases (DARC) from racially biased disinformation (RBD): the racial case is invoked as context, not as the site of the harm itself. These claims deployed the same operations (SI, AR) but directed them against white actors' reputations, not against Marielle or Black communities as primary targets. Without this distinction, a researcher studying disinformation around the Marielle case might count false claims about white political actors as racial

disinformation, when they are political disputes that use a racial case as backdrop. This would overcount racial disinformation and obscure what makes it structurally distinct from disinformation that directly harms Black individuals and communities. Both ecosystems are documented in the replication archive; only RBD constitutes the analytical corpus.

***Table 1**. Corpus overview by source (2018–2023).*

Source	Total collected	Racially filtered	Confirmed RBD	Inclusion rate	Content
Agência Lupa	453	143 (31.6%)	27	18.9%	Full text
Aos Fatos	3,221	49 (1.5%)	7	14.3%	Title only*
Total	**3,674**	**192 (5.2%)**	**34**	**17.7%**	**—**

Note: Aos Fatos renders content via JavaScript; title only available. RBD = racially biased disinformation.

Discussion and future directions

The four findings reported here converge on a structural argument: racially biased disinformation in Brazil is a structured practice whose operations are explicable by the racial inequalities on which each depends. *Identity substitution* (SI) works because peripheral Black individuals have low public recognition. *Criminalization* (CR) works because punitivist narratives are hegemonic. *Relational association* (AR) works because racial-territorial coding equates favela with criminality. *Moral defamation* (DM) works, particularly against Black women politicians, because intersecting racial and gender hierarchies construct certain bodies as politically illegitimate. *Decontextualization* (DC) works because affective media logics prioritize emotional response over deliberative assessment. Each operation exploits a distinct dimension of racial inequality; their combination exploits multiple dimensions simultaneously, producing what we term operational complexity, a structural asymmetry between disinformation's harm and refutation's reach that our data suggest but do not directly demonstrate. Fact checkers, therefore, need operation-aware refutation strategies, and platforms need detection systems calibrated to compound rather than single-operation claims (Freelon & Wells, 2020).

The lower *criminalization* (CR) inter-coder agreement (κ = 0.10, 53.3%) carries analytical weight beyond its methodological implications. Criminalization is the most frequently deployed operation (76.5%) and simultaneously the most difficult to detect from fact-checker text alone because it operates significantly through the implicit and unstated. A fact-checker refuting the claim that a police victim "was a trafficker" does not need to reproduce the original attribution to debunk it, meaning the debunking text lacks the linguistic markers that automated or human coding would expect. A disinformation operation that functions through what it does not need to say is precisely the most deeply embedded in social structure. This is the digital expression of what Almeida (2018) calls structural racism, a form of violence that does not require articulation because it is already assumed by those who receive and share the claim.

The recirculation of disinformation about Marielle Franco two years after her assassination illustrates a broader principle. Following Pollak (1989), racially biased disinformation functions as sustained memory contestation: it does not resolve when debunked but reconstitutes in each new news cycle that brings its target back into public attention. Platform affordances enabling content to resurface are not neutral; they actively maintain conditions for counter-memory production against Black victims of racial violence. Addressing operational complexity is therefore not only a research agenda but a platform design imperative, and the framework proposed here offers a replicable starting point where three conditions converge: racialized stratification producing peripheral communities with low public visibility, a punitivist media culture normalizing criminal attributions to racial out-groups, and high social media penetration enabling compound claim distribution.

Translating this framework into practice involves concrete steps from three distinct actors. For platforms, the finding that Facebook accounts for 73.5% of confirmed cases identifies where intervention

is most urgent. Content moderation systems could use the five operation codes to flag recurring patterns of race-based disinformation before they spread further, prioritizing content that combines two or more operations, since compound claims account for 72% of confirmed cases and are the hardest to neutralize through single-operation responses. The framework's pattern-based logic is also well suited to inform computational models and train algorithms to recognize forms of racial disinformation that current automated systems overlook because they operate through culturally embedded stereotypes rather than explicit slurs. For fact-checking teams, checking each claim about a Black victim of violence against all five operations would catch secondary ones that might otherwise persist after the primary claim is addressed. For educators and media-literacy programs, naming the SI+CR and DC+CR patterns as recognizable templates offers a heuristic more actionable than general warnings about disinformation, one that can be taught and applied in real time.

Methods

Research design and data sources

This study employs a qualitative research design with descriptive quantitative components (Creswell & Clark, 2017). Following an abductive qualitative content analysis approach, cases move iteratively between the empirical corpus and the five-operation typology: the typology was developed inductively from an initial close reading of a subset of confirmed cases, then applied deductively to the full candidate set, with codes revised when cases resisted the existing categories. Data were collected from two IFCN-certified agencies (the International Fact-Checking Network, a global consortium hosted by the Poynter Institute that vets fact-checking organizations against a code of principles covering methodological transparency, non-partisanship, and open corrections): *Agência Lupa* (agencialupa.org) and *Aos Fatos* (aosfatos.org). A third agency (*Agência Pública*) was excluded because its editorial model is long-form investigative journalism rather than standardized claim-by-claim verification; items without explicit verdicts cannot be reliably coded as confirmed disinformation. Data from *Agência Lupa* were collected via the WordPress REST API using nine keyword searches; data from *Aos Fatos* via RSS feed and tag-based filtering. A technical constraint shaped the *Aos Fatos* collection: the site renders article body content via client-side JavaScript, making full-text extraction unavailable through static methods, so *Aos Fatos* items were coded from titles only. The period runs from January 2018, the year of Marielle Franco's assassination, to December 2023. The complete racial filter dictionary is provided in Appendix A.

Corpus construction and inclusion criteria

The combined corpus comprised 3,674 fact-checks. A dictionary-based racial filter identified 192 candidate items (a keyword search, listed in full in Appendix A, that flags any fact-check whose title or available text contains terms signaling racial identity, racially coded territory such as "favela," police-violence vocabulary, documented cases of racially targeted violence, or explicit references to structural racism, so that items can then be manually reviewed for relevance). Manual coding applied a two-criteria inclusion rule: an item was confirmed as RBD only when (a) a specific Black individual or racially coded group was the primary target, and (b) at least one discursive operation was deployed against that target. Both criteria were required simultaneously. Items meeting only one criterion were classified as DARC when the primary operation targeted a white actor using a racial event as context.

Discursive operations framework

***Table 2**. Discursive operations framework for racially biased disinformation (Brazil, 2018–2023).*

Code	Operation	Definition	Mechanism	*N* (%)
SI	Identity substitution	False image or identity attributed to a Black victim to misrepresent them as a criminal actor	Exploits low recognition of peripheral individuals; visual immediacy makes false attribution plausible	17 (50.0%)
CR	Criminalization	Fabricated or decontextualized attribution of criminal history or behavior to a Black individual or group	Inverts victim/perpetrator relationship; State violence retroactively framed as legitimate response	26 (76.5%)
DM	Moral defamation	Attack on the victim's private life, sexuality, or relationships through false or misleading claims	Constructs victim as undeserving of solidarity rather than as a threat (cf. CR); produces silencing	5 (14.7%)
AR	Relational association	Delegitimization through fabricated or exaggerated associations with criminal actors or racially coded territories	Exploits racial-territorial coding (favela = crime); difficult to refute without validating the premise	6 (17.6%)
DC	Decontextualization	Factually accurate content placed in a misleading frame to criminalize Black movements or communities	Not technically false; resists debunking; persists longer in circulation	14 (41.2%)

Note: Operations are not mutually exclusive; 72% of confirmed cases exhibit two or more simultaneous operations. Compound cases coded with "+" (co-primary) or ">" (primary > secondary).

The five operations are not mutually exclusive, and their co-presence is analytically significant. Compound cases are coded with all operations present. Where one is clearly dominant, it is indicated with ">" notation (e.g., CR > DM = criminalization primary, moral defamation secondary); co-primary operations are indicated with "+" (e.g., CR + SI = both equally central to the claim's delegitimizing function).

Inter-coder reliability

Cohen's kappa is a statistic that measures how much two independent coders agree beyond what would be expected by chance alone, ranging from 0 (no better than chance) to 1 (perfect agreement); values above 0.60 are generally considered substantial (Landis & Koch, 1977). A second coder with expertise in Brazilian politics independently coded a stratified random sample of 40 items (20 RBD, 20 excluded; 20.8% of the manually reviewed subset). Both authors are Brazilian political scientists with sustained engagement with the media environment and racial politics analyzed in this study; this shared background informed the conceptual development of the framework and the interpretation of borderline coding decisions. Cohen's kappa for the primary inclusion variable was $\kappa = 0.70$ ($p < .001$, $N = 40$), indicating substantial agreement (Landis & Koch, 1977), with 85.0% agreement. For discursive operations, kappa was calculated per operation individually for the 15 items both coders confirmed as RBD: AR ($\kappa = 1.00$), DM ($\kappa = 0.60$), DC ($\kappa = 0.40$), SI ($\kappa = 0.32$), and CR ($\kappa = 0.10$). The lower agreement on CR is discussed in the following section as an analytical finding itself. Six discordances on the inclusion variable were resolved

by consensus. Full inter-coder reliability data, including all coding decisions and resolved discordances, are deposited at Harvard Dataverse (see Data Availability).

***Table 3**. Inter-coder reliability: Cohen's kappa by variable.*

Variable	κ	% Agreement	N	Interpretation
CONFIRMAR — inclusion decision	0.70	85.0%	40	Substantial
AR — Relational association	1.00	100.0%	15	Perfect
DM — Moral defamation	0.60	86.7%	15	Substantial
DC — Decontextualization	0.40	73.3%	15	Moderate
SI — Identity substitution	0.32	66.7%	15	Fair
CR — Criminalization	0.10	53.3%	15	Slight

Note. κ for CONFIRMAR: N = 40 stratified items (20 SIM, 20 NÃO; 20.8% of manually reviewed subset). Operation-level κ: N = 15 items (both coders confirmed SIM). Interpretation: Landis & Koch (1977). Lower CR agreement reflects title-only content access for Aos Fatos items; see Limitations.

Limitations

Four limitations bear noting. (1) Fact-checker archives systematically underrepresent encrypted content on WhatsApp, the dominant channel for disinformation in Brazil (Melo et al., 2019), while overrepresenting high-visibility Facebook claims (73.5% of confirmed cases; Figure 3); our findings therefore capture the visible surface of a larger phenomenon. (2) *Aos Fatos* title-only coding limits operation-level classification for 7 of 34 confirmed cases, contributing to the lower inter-coder agreement on CR in that subset. Future research could overcome this constraint by using headless-browser scraping tools (e.g., Selenium or Playwright) that execute client-side JavaScript before extraction, or by requesting bulk data or API access directly from *Aos Fatos*, either of which would enable full-text operation-level coding of that agency's fact-checks. (3) Seven cases (20.6%) involve a single individual, meaning findings about the Marielle Franco cluster should be read as case-specific as well as pattern-revealing. (4) Framework generalizability to other high-inequality democracies requires empirical validation through comparative research, which we identify as a priority for future work.

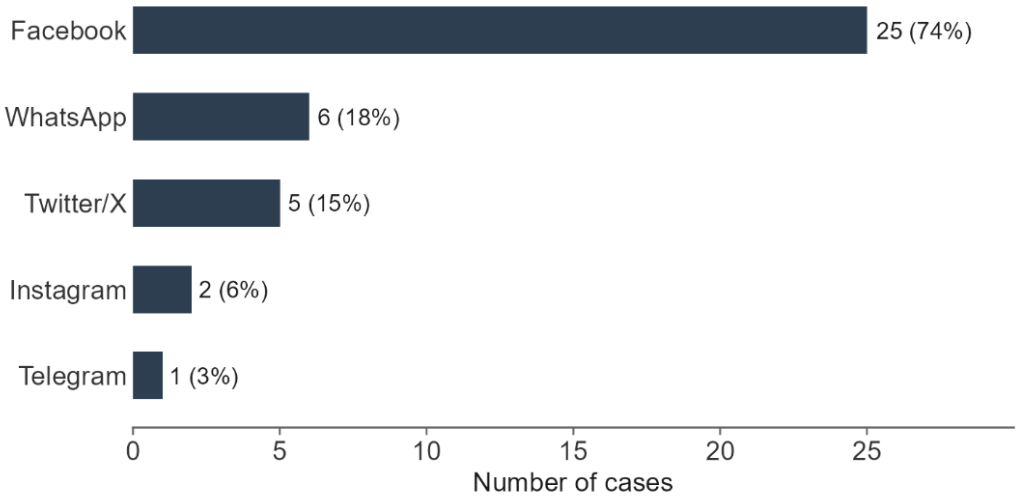


Figure 3. Primary platforms of disinformation circulation in confirmed cases (N = 34; some cases cite multiple platforms).
Source: authors' coding of fact-check article descriptions, 2018–2023.

Bibliography

Almeida, S. (2018). *O que é racismo estrutural* [What is structural racism]? Letramento.

Anastácio, K. (2024). Framing disinformation through legislation: Evidence from policy proposals in Brazil. *Harvard Kennedy School (HKS) Misinformation Review*, *5*(3). https://doi.org/10.37016/mr-2020-150

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, *1*, 13–29. https://doi.org/10.1038/s44159-021-00006-y

Fórum Brasileiro de Segurança Pública. (2023). *Anuário brasileiro de segurança pública 2023* [Brazilian yearbook of public cecurity 2023]. https://forumseguranca.org.br/anuario-brasileiro-seguranca-publica/

Freelon, D., McIlwain, C., & Clark, M. (2016). *Beyond the hashtags: #Ferguson, #Blacklivesmatter, and the online struggle for offline justice*. Center for Media & Social Impact. https://cmsimpact.org/resource/beyond-hashtags/

Freelon, D., & Wells, C. (2020). Disinformation as political communication. *Political Communication*, *37*(2), 145–156. https://doi.org/10.1080/10584609.2020.1723755

Gehrke, M., & Amit-Danhi, E. R. (2025). Gendered disinformation as violence: A new analytical agenda. *Harvard Kennedy School (HKS) Misinformation Review*, *6*(3). https://doi.org/10.37016/mr-2020-177

IBGE. (2023). *Censo Demográfico 2022: Identificação étnico-racial da população, por sexo e idade: Resultados do universo* [Demographic census 2022: Ethnic-racial identification of the population, by sex and age: Universe results]. Instituto Brasileiro de Geografia e Estatística. https://biblioteca.ibge.gov.br/visualizacao/periodicos/3105/cd_2022_etnico_racial.pdf

INFOPEN. (2021). *Levantamento nacional de informações penitenciárias* [National prison information survey]. Ministério da Justiça. https://www.gov.br/senappen/pt-br/servicos/sisdepen

Kuo, R., & Marwick, A. (2021). Critical disinformation studies: History, power, and politics. *Harvard Kennedy School (HKS) Misinformation Review*, *2*(4). https://doi.org/10.37016/mr-2020-76

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Mbembe, A. (2018). *Necropolitics*. Duke University Press. https://doi.org/10.1215/9781478007227

Media Manipulation Casebook. (2021). *Racialized disinformation*. Harvard Kennedy School: Shorenstein Center on Media, Politics and Public Policy. https://mediamanipulation.org/definitions/racialized-disinformation/

Melo, P., Messias, J., Resende, G., Garimella, K., Almeida, J. M., & Benevenuto, F. (2019). WhatsApp monitor: A fact-checking system for WhatsApp. *Proceedings of the International AAAI Conference on Web and Social Media*, *13*, 676–677. https://doi.org/10.1609/icwsm.v13i01.3271

Moura, C. (1994). *Dialética radical do Brasil negro* [Radical dialectics of Black Brazil]. Editora Anita.

Nunes Junior, E. M. (2021). Mídia, fake news e racismo: O punitivismo dos boatos como legitimador da violência [Media, fake news, and racism: Punitivism in rumours as a legitimizer of violence]. *Revista Brasileira de Segurança Pública*, *15*(1), 10–21. https://doi.org/10.31060/rbsp.2021.v15.n1.1122

Omi, M., & Winant, H. (1994). *Racial formation in the United States* (2nd ed.). Routledge.

Pollak, M. (1989). Memória, esquecimento, silêncio [Memory, forgetting, silence]. *Estudos Históricos*, *2*(3), 3–15. https://periodicos.fgv.br/reh/article/view/2278

Reddi, M., Kuo, R., & Kreiss, D. (2021). Identity propaganda: Racial narratives and disinformation. *New Media & Society*, *23*(8), 2161–2173. https://doi.org/10.1177/14614448211029293

Romão, D. M. M. (2013). *Jornalismo policial: Indústria cultural e violência* [Police journalism: Cultural industry and violence] [Master's thesis, Universidade de São Paulo]. https://repositorio.usp.br/item/002400807.

Sodré, M. (2006). *As estratégias sensíveis: Afeto, mídia e política* [Sensitive strategies: Affect, media, and politics]. Vozes.

Vargas, J. H. C. (2004). Hyperconsciousness of race and its negation: The dialectic of white supremacy in Brazil. *Identities*, *11*(4), 443–470. https://doi.org/10.1080/10702890490883803

Wacquant, L. (2007). *Punishing the poor: The neoliberal government of social insecurity*. Duke University Press.

Funding
No external funding was received for this research.

Competing interests
The authors declare no competing interests.

Ethics
This study uses exclusively publicly available journalistic content published by IFCN-certified organizations. No human subjects research was conducted. No personal data beyond what is available in published fact-check articles is reported.

Copyright
This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are properly credited.

Data availability
All materials needed to replicate this study are available via the Harvard Dataverse: https://doi.org/10.7910/DVN/CK1IO1

Appendix: Racial dictionary: Filter terms

The following terms were used to filter the full corpus of 3,674 fact-checks for racial relevance. Terms are organized across five semantic dimensions. The filter was applied to title and available content text. Items matching at least one term in any dimension were retained for manual review (N = 192).

Dimension 1: Identity terms
negro, negra, negros, negras, pardo, parda, pardos, pardas, preto, preta, afrodescendente, afro-brasileiro, afro-brasileira, quilombola.

Dimension 2: Racially coded territorial terms
favela, favelado, favelada, comunidade, periférico, periférica, periferia, morro, complexo do, subúrbio, baixada. Note: "Favela" and related terms function as racial proxies in Brazilian public discourse, where peripheral territory is consistently associated with Black and mixed-race populations (Vargas, 2004). Items using these terms were retained even without explicit racial identity language.

Dimension 3: Police violence vocabulary
auto de resistência, resistência seguida de morte, bala perdida, operação policial, intervenção policial, caveirão, UPP, BOPE, chacina, execução, homicídio policial, violência policial.

Dimension 4: Documented cases of racially targeted violence
marielle, marielle franco, agatha felix, ágatha, joão pedro, joão pedro matos, genivaldo, genivaldo jesus, beto freitas, marcus vinicius, marcos vinicius, douglas rafael, DG esquenta, eduardo jesus, jonatha dalber, maria eduarda, acari.

Dimension 5: Structural racism terms
racismo, racial, racismo estrutural, racismo institucional, discriminação racial, genocídio negro, encarceramento em massa, criminalização da pobreza, desigualdade racial, população negra.