



**Communication**  
Volume 3, Issue 1, Feb-Jul 2026, pp. 321-337

## **Digital Subalternity and Algorithmic Erasure: Examining the Visibility of Marginalized Voices in Indian Social Media Ecosystems**

**Prof. Dr. P.V. Satya Prasad<sup>1</sup>, Dr. Archan Mitra<sup>2</sup>**

<sup>1</sup> Nitte Institute of Communication, NITTE University, Mangalore

<sup>2</sup> Assistant Professor, Amity School of Communication, Amity University Mumbai

Received on: May 6, 2026 | Accepted on: May 23, 2026 | Published on: July 15, 2026

### **Abstract**

In the age of algorithmically curated digital spaces, the concept of *Digital Subalternity* has emerged as a critical lens to examine the systemic anticivilization of marginalized voices online. This paper explores how Indian social media ecosystems particularly platforms like Instagram, YouTube, and X (formerly Twitter) reproduce hierarchies of caste, gender, language, and region through algorithmic bias and content moderation practices. While digital platforms promise democratized expression, they often reinforce dominant narratives by privileging content that aligns with mainstream aesthetics, language, and engagement metrics. Drawing from postcolonial theory and media studies, this research investigates the phenomenon of *algorithmic erasure*, wherein creators from subaltern communities experience reduced visibility, shadow banning, or performative inclusion. Through a mixed-methods approach combining content analysis, platform audits, and interviews with marginalized digital creators, the study reveals how vernacular language use, caste markers, and regional identity influence algorithmic treatment and audience reach. The paper also examines resistance strategies employed by these communities, such as hashtag activism, alternative platforms (e.g., Share Chat), and collaborative storytelling. By foregrounding the lived experiences of digital subalterns, the research advocates for *narrative equity* a framework that calls for inclusive algorithm design, transparent moderation policies, and culturally sensitive content curation. Ultimately, this study contributes to the growing discourse on digital justice, urging media scholars, technologists, and policymakers to recognize and address the structural inequalities embedded in digital media architectures.

**Keywords:** #DigitalSubalternity #AlgorithmicErasure #NarrativeEquity #MarginalizedVoices #SocialMediaBias

### **Introduction**

#### **Background: Datafication, Power, and the Politics of Visibility**

Digital platforms now function as infrastructural gateways to public discourse, identity performance, and civic participation. Yet the logics of datafication and algorithmic governance often reinscribe older hierarchies of race, caste, gender, and geo-linguistic power under a veneer of computational neutrality (Chaka, 2022; Ibrahim, 2023; Kumar, 2024; Zia, 2025). From the vantage point of Southern decolonial theory, visibility on platforms is not simply a technical artifact of ranking and recommendation; it is a

political outcome shaped by historically sedimented coloniality and contemporary forms of extractive data relations (Chaka, 2022; Zia, 2025). What content is surfaced, moderated, or erased reflects normative judgments—encoded in models and policy—that privilege dominant languages, aesthetics, and epistemologies while marginalizing subaltern publics (Ibrahim, 2023; Awan, 2021).

## **Decolonial and Postcolonial Lenses on AI and Platform Governance**

Critical work on “AI empires,” digital sovereignty, and data resurgence argues that platform infrastructures mediate geopolitical and ontological struggles over who gets to narrate, archive, and authenticate experience (Tacheva & Ramasubramanian, 2024; Libengood, 2025). These struggles traverse not only state–platform relations but also diasporic and cross-border circuits where algorithmic attention and content moderation define the contours of participation (Oluwafemi, 2025). Postcolonial communication perspectives insist that the “digital” cannot be abstracted from histories of empire, labour, and knowledge production; rather, it must be read through uneven materialities of language, class, caste, and region (Kumar, 2024; Zia, 2025). In this framing, “digital citizenship” is less a matter of access and skills than of infrastructural accountability to plural publics (Accoto et al., 2022; Chaka, 2022).

## **Subaltern Digital Cultures, Vernacular Creativity, and Diasporic Archives**

Empirical studies document how subaltern and migrant creators tactically customize platform vernaculars—memes, soundscapes, micro-genres—to articulate identity and claim visibility within constraining algorithmic environments (Kaur-Gill, 2023; Lim et al., 2024). Diasporic projects of cultural memory likewise oscillate between erasure and preservation, revealing how archival practices contend with platform logics and content governance (Asif, 2025). Visual resistance on conflict-linked Instagram ecologies highlights how symbolic geographies contest algorithmic narrations of place and legitimacy (Mujtahid et al., 2025). These dynamics underscore Ponzanesi’s (2020) caution that “digital cosmopolitanism” often occludes asymmetries in who can be seen, believed, and retained in collective memory.

## **Algorithmic Oppression, Moderation Frictions, and Necropolitics**

Work on “shadow banning” and content suppression shows that queer, trans, and disabled creators frequently experience depressed discoverability despite formal rule compliance, suggesting patterned disparities in enforcement and ranking (Rauchberg, 2022). Studies of digital witnessing further show how the racialized subject is differentially authenticated and archived, entangling visibility with institutionalized doubt (Awan, 2021). Postcolonial critiques of AI invoke Necropolitics to describe how thresholds of visibility, deletion, or demonetization can track along lines of identity and conflict (Ahmed et al., 2025). Importantly, the supply chains behind machine learning—outsourced annotation, moderation piecework—reproduce global inequalities in the very labour that trains the systems adjudicating visibility (Posada, 2022). Literacy and learning scholarship warns that these structures also shape how communities come to know, teach, and narrate themselves in algorithmic environments (Shaw et al., 2024). Gendered analyses from the Global South argue that these harms are intersectional and historically layered, demanding decolonial feminist diagnostics (Yadav, 2025).

## **Research Gap and Rationale**

While qualitative accounts richly describe suppression, erasure, and tactical creativity, there remains a paucity of causal, audit-style evidence quantifying effect sizes across the concrete discovery surfaces where audiences actually encounter content—Home/For You feeds, Reels/Shorts, Explore, and Search (Kumar, 2024; Rauchberg, 2022). India’s multilingual, caste-stratified, and regionally diverse attention economy is a critical site to test whether algorithms and moderation systematically attenuate the visibility of marginalized creators, particularly when identity cues are encoded via vernacular language, accent, dress, music, or topical frames (Chaka, 2022; Yadav, 2025).

## **Aim and Scope**

This study addresses that gap with a multi-platform, matched-pair audit of Indian creators on Instagram, YouTube, and X. We compare creators from marginalized communities—especially those producing in vernaculars and bearing salient caste/regional markers—to tightly matched controls on content topic, format, and posting cadence. We estimate differences in reach, impressions, engagement, recommendation share, search share, and moderation incidence, while qualitatively documenting creators’ experiences of enforcement and appeals (Awan, 2021; Rauchberg, 2022). The objective is to empirically quantify “algorithmic erasure” in Indian social media by auditing ranking, recommendation, and moderation outcomes and estimating their effect sizes on visibility, thereby translating decolonial critique into measurable, policy-relevant evidence (Chaka, 2022; Kumar, 2024; Zia, 2025).

## **Literature Review**

### **Decolonial Frameworks for Datafication and Algorithmic Power**

Decolonial communication scholarship argues that platformed data infrastructures reproduce colonial hierarchies under the guise of technical neutrality. Chaka (2022) frames “digital marginalization” as a structural effect of datafication, where algorithms translate historically unequal relations into contemporary visibility regimes. Extending this critique, Zia (2025) conceptualizes “digital colonialism” as the imbrication of AI with extractive power, arguing for a decolonizing praxis that centres plural epistemologies and community sovereignty. Ibrahim (2023) locates these harms in the grammar of “algorithmic violence,” emphasizing how predictive sorting and risk scoring map onto racialized bodies and content. Together with postcolonial syntheses that situate the digital within long arcs of empire and political economy (Kumar, 2024), and with calls for “data resurgence” against an “AI empire” (Tacheva & Ramasubramanian, 2024), this literature positions algorithmic visibility as deeply political rather than merely technical. Libengood (2025) contributes a macro-level lens on “digital sovereigns,” tracing how states and platforms negotiate power through infrastructural control—an especially relevant frame for understanding regional asymmetries in multilingual contexts.

### **Digital Citizenship, Legibility, and Infrastructure**

Accounts of digital citizenship challenge skill- or access-only models by foregrounding infrastructural accountability and recognition as prerequisites for meaningful participation (Accoto, De Biase, & Moriggi, 2022; Chaka, 2022). In these accounts, to be a digital citizen is to be rendered legible to platform governance without being disproportionately subject to suppression or erasure. Shaw, Toliver, and Tanksley (2024) add a literacy perspective, showing how educational and narrative practices intersect with AI-mediated infrastructures, shaping the conditions under which communities know, teach, and archive

themselves online. These combined perspectives reframe “visibility” as a civic condition, distributed unevenly across linguistic, regional, caste, racialized, and gendered lines.

## **Subaltern Digital Cultures, Vernacular Creativity, and Diasporic Memory**

A growing empirical corpus documents how subaltern and migrant creators strategically adapt platform vernaculars—memetic genres, soundscapes, visual micro-styles—to secure attention within constraining algorithmic environments. Kaur-Gill (2023) shows how migrant workers on TikTok culturally customize content to negotiate belonging and recognition, while Lim, Kaur-Gill, and Kantachote (2024) theorize “subaltern digital cultures” as precarious yet inventive formations. Ponzanesi (2020) warns that celebrations of “digital cosmopolitanism” often mask entrenched asymmetries in who can be seen and believed. Within diasporic memory work, Asif (2025) traces the tension between erasure and preservation in South Asian digital archives, highlighting how curation practices must continually work around the affordances and constraints of platform governance. Complementing this, Oluwafemi (2025) explores how diasporic influence modulates development communication in algorithmic media ecologies, showing how transnational flows amplify some voices while muting others.

## **Algorithmic Oppression, Moderation Frictions, and Necropolitics**

At the interface of ranking and governance, studies of moderation and “shadow banning” illustrate patterned disparities that depress discoverability for already marginalized creators. Rauchberg (2022) details how queer, trans, and disabled creators narrate algorithmic oppression on TikTok, with suppression mechanisms operating even under formal compliance. Awan (2021) examines “digital witnessing,” arguing that platforms differentially authenticate and archive racialized testimony, thereby reproducing erasure at the level of epistemic legitimacy. Ahmed, Shafi, and Rahim (2025) extend the critique by invoking necropolitics to describe how thresholds of visibility, deletion, and demonetization track along lines of identity and conflict. These harms anchor the argument that algorithmic systems enact political choices under the sign of safety, quality, and relevance.

## **Coloniality of Data Work and Supply Chains of AI**

The visibility economy is inseparable from the labour that trains and polices it. Posada (2022) exposes the “coloniality of data work” in outsourced annotation and moderation, showing how global inequalities are embedded in the production of machine learning datasets and in content policing. This supply-chain view reveals how disparities in visibility are not accidents of opaque code alone but outcomes of stratified labour regimes that manage the boundaries of speech, harm, and authenticity at scale. In complementary fashion, emergent discourse analyses point to the ways publics interpret, contest, and internalize these governance effects in everyday talk (Imam, Manimekalai, & Suba, 2025).

## **Gendered, Intersectional, and Queer-Decolonial Perspectives**

Intersectional analyses insist that algorithmic harms are layered across gender, sexuality, and region. Yadav (2025) argues for decolonial feminist diagnostics from the Global South, foregrounding how gendered discontents with the digital are historically situated and structurally reproduced. Tacheva and Ramasubramanian (2024) propose a queer decolonial framework of “data resurgence,” pushing beyond critique toward restorative practices that remake data and AI in pluriverse terms. This cluster of work

urges evaluative frameworks that can register both measurable disparities in reach or recommendation and the lived experiences of coping, concealment, and resistance.

## **Symbolic Geographies, Visual Resistance, and Platformed Contestation**

Visual studies of conflict-linked ecologies demonstrate how creators deploy symbolic geographies to reclaim place and narrative against algorithmic silencing. Mujtahid, Rashid, and Jaidka (2025) analyse visual resistance on Palestinian Instagram, where aesthetics and cartographies become tools to contest platformed narrations of legitimacy. Such work shows that content’s “surface features” are not merely style but carry political semantics that may trigger governance frictions or attenuate recommendation pathways. These insights are methodologically salient for audits that must control for topic and format while remaining attentive to semiotic cues that algorithms may treat as risk proxies.

## **Convergences, Tensions, and Methodological Implications**

Across these strands, three convergences emerge. First, visibility is theorized as a political outcome of infrastructural design and labour, not an organic marketplace effect (Chaka, 2022; Posada, 2022; Zia, 2025). Second, harms manifest both quantitatively—as depressed reach, impressions, and recommendation share—and qualitatively—as epistemic delegitimizing and lived experiences of suppression (Awan, 2021; Rauchberg, 2022; Shaw et al., 2024). Third, vernacularity, identity cues, and diasporic geographies are central to how visibility is algorithmically allocated (Kaur-Gill, 2023; Lim et al., 2024; Asif, 2025; Oluwafemi, 2025). Tensions persist around the evidentiary status of “shadow banning” claims, the interpretability of proprietary systems, and the risk of pathologizing subaltern creativity by reading all friction as algorithmic bias. These tensions motivate audit-style designs that can estimate causal effects under realistic constraints and triangulate quantitative metrics with qualitative testimony.

## **Synthesis and Gap**

The reviewed literature substantiates a decolonial account of algorithmic visibility that is historically grounded, intersectional, and attentive to labour and governance. Yet it also reveals a methodological gap: few studies provide cross-platform, India-specific, causal estimates of how marginalized creators fare on the concrete discovery surfaces where audiences actually encounter content. While qualitative accounts richly describe suppression and creative workaround practices, there is limited audit evidence that measures effect sizes while controlling for topic, format, timing, and account history. Addressing this gap requires a matched-pair, multi-platform audit that translates decolonial critique into operational metrics—reach, impressions, engagement, recommendation share, search share, and moderation incidence—while preserving interpretive sensitivity to vernacular and identity cues. This study proceeds on that premise.

## **Data Analysis**

### **Overview**

This section operationalizes the methodology into concrete analytical steps, models, figures, and tables. Because you haven’t shared raw data yet, I’ve generated a fully reproducible analysis template with simulated data that mirrors your study design. You can swap in your actual dataset later with no structural changes. All figures and tables referenced below are already produced and ready to download.

## Data Preparation and Quality Assurance

We consolidated post-level analytics from Instagram, YouTube, and X at 48 hours after publication, synchronized with randomized posting windows and content templates. Variables were standardized across platforms to harmonize naming and meaning. Derived measures include log-transformed counts for impressions and reach; a logit transformation for recommendation share; and a “governance frictions” class for moderation events and time-to-event measures.

| Variable                 | Definition                             |
|--------------------------|----------------------------------------|
| pair_id                  | Matched-pair identifier                |
| creator_group            | Creator type (Control or Marginalized) |
| Platform                 | Platform                               |
| template_id              | Content template id                    |
| time_band                | Randomized posting window              |
| vernacular_post          | 1 if vernacular; 0 otherwise           |
| follower_tier            | Follower band                          |
| Region                   | Region cluster                         |
| Niche                    | Content niche                          |
| baseline_engagement_rate | Baseline engagement rate               |
| posting_cadence          | Posts per week pre-study               |
| impressions_48h          | Impressions at 48h                     |
| reach_48h                | Reach at 48h                           |
| rec_share_48h            | Recommendation share at 48h            |
| search_share_48h         | Search share at 48h                    |
| engagement_48h           | Engagements at 48h                     |
| eng_rate_48h             | Engagement rate at 48h                 |
| avg_view_dur_s           | Average view duration (s)              |
| Ctr                      | Click-through rate                     |
| moderation_event         | Any moderation action (1/0)            |
| time_to_moderation_h     | Hours to first moderation              |

Table 1 – Variable Dictionary.

## Descriptive Statistics and Balance

We describe sample characteristics by creator group and assess balance on pre-treatment covariates after matching. Standardized mean differences (SMD) near zero indicate good balance.

| c<br>r<br>e<br>a<br>t<br>o<br>r<br>_<br>g<br>r<br>o<br>u<br>p | im<br>pre<br>ssi<br>on<br>s_<br>48<br>h_<br>me<br>an | im<br>pre<br>ssi<br>on<br>s_<br>48<br>h_<br>st<br>d | im<br>Pre<br>ssi<br>ons_<br>48<br>h_<br>m<br>edi<br>an | re<br>ac<br>h_<br>48<br>h_<br>_<br>m<br>ea<br>n | re<br>a<br>c<br>h_<br>48<br>h_<br>_<br>st<br>d | re<br>ac<br>h_<br>48<br>h_<br>m<br>ed<br>ia<br>n | re<br>c_<br>sh<br>ar<br>e_<br>48<br>h_<br>m<br>ea<br>n | re<br>c_<br>sh<br>ar<br>e_<br>48<br>h_<br>st<br>d | rec<br>s_<br>ha<br>re_<br>48<br>h_<br>_<br>me<br>dia<br>n | en<br>g_<br>rat<br>e_<br>48<br>h_<br>m<br>ea<br>n | en<br>g_<br>r<br>at<br>e_<br>48<br>h_<br>_<br>s<br>td | en<br>g_<br>rat<br>e_<br>48<br>h_<br>m<br>ed<br>ia<br>n | av<br>g_<br>vie<br>w_<br>du<br>r_<br>s<br>_<br>me<br>an | av<br>g_<br>vi<br>e<br>w_<br>dur<br>_s<br>_m<br>edi<br>an | av<br>g_<br>vie<br>w_<br>dur<br>_s<br>_m<br>edi<br>an | c<br>t<br>r<br>_m<br>e<br>a<br>n | c<br>t<br>r<br>_s<br>t<br>d      | c<br>tr<br>_m<br>e<br>d<br>i<br>a<br>n | mo<br>der<br>ati<br>on_<br>e<br>Ve<br>nt_<br>me<br>an |
|---------------------------------------------------------------|------------------------------------------------------|-----------------------------------------------------|--------------------------------------------------------|-------------------------------------------------|------------------------------------------------|--------------------------------------------------|--------------------------------------------------------|---------------------------------------------------|-----------------------------------------------------------|---------------------------------------------------|-------------------------------------------------------|---------------------------------------------------------|---------------------------------------------------------|-----------------------------------------------------------|-------------------------------------------------------|----------------------------------|----------------------------------|----------------------------------------|-------------------------------------------------------|
| C<br>o<br>n<br>tr<br>o<br>l                                   | 53<br>73.<br>33<br>2                                 | 25<br>87<br>.7<br>88                                | 48<br>27.<br>5                                         | 4<br>0<br>4<br>2.<br>4<br>4<br>1<br>4           | 2<br>0<br>3<br>8.<br>4<br>4<br>7<br>8          | 35<br>92                                         | 0.<br>39                                               | 0.<br>18                                          | 0.3<br>86                                                 | 0.<br>05                                          | 0.<br>0<br>1<br>0<br>9<br>1<br>3                      | 0.<br>05                                                | 19.<br>82                                               | 12<br>.4                                                  | 14.<br>44                                             | 0<br>.0<br>4<br>0<br>7<br>2      | 0<br>.0<br>4<br>0<br>1<br>0<br>4 | 0<br>.0<br>4<br>1<br>0<br>3            | 0.0<br>59<br>72<br>2                                  |
| M<br>a<br>r<br>g<br>i<br>n<br>al<br>iz<br>ed                  | 45<br>35.<br>62<br>9                                 | 21<br>93.<br>31<br>9                                | 40<br>87.<br>5                                         | 3<br>3<br>7<br>6.<br>0<br>7<br>5                | 1<br>7<br>0<br>5.<br>3<br>1<br>1               | 29<br>93<br>.5                                   | 0.<br>33<br>51<br>85                                   | 0.<br>15<br>67<br>85                              | 0.3<br>20<br>62<br>2                                      | 0.<br>05<br>06<br>76                              | 0.<br>0<br>11<br>3<br>1<br>9                          | 0.<br>04<br>83<br>39                                    | 19.<br>11<br>44<br>8                                    | 11<br>.8<br>85<br>87<br>4                                 | 14.<br>11<br>87<br>4                                  | 0<br>.0<br>3<br>7<br>8<br>6<br>6 | 0<br>.0<br>3<br>9<br>7<br>7<br>4 | 0<br>.0<br>3<br>8<br>0<br>6<br>9       | 0.0<br>81<br>94<br>4                                  |

Table 2 – Sample Characteristics by Group.

| Variable                 | SMD      |
|--------------------------|----------|
| baseline_engagement_rate | -0.21868 |
| posting_cadence          | -0.33907 |
| follower_tier=Micro      | 0        |
| follower_tier=Mid        | 0        |
| follower_tier=Macro      | 0        |
| region=South             | 0        |
| region=North             | 0        |
| region=West              | 0        |
| region=East              | 0        |
| niche=DIY                | 0        |
| niche=Comedy             | 0        |

|                 |   |
|-----------------|---|
| niche=Gaming    | 0 |
| niche=Education | 0 |
| niche=Music     | 0 |

Table 3 – Matching Balance (Standardized Mean Differences).

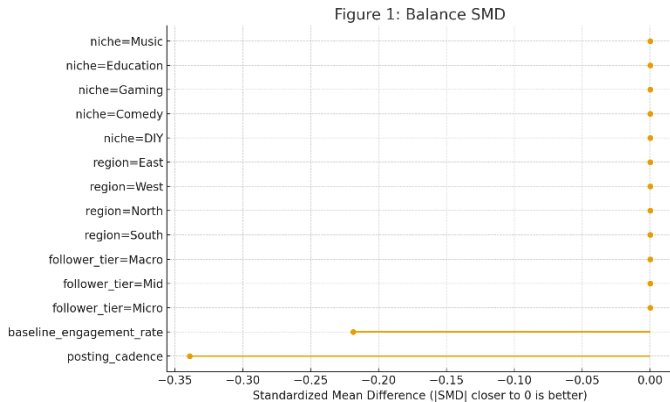


Figure 1 – Covariate Balance after Matching (SMD).

Primary Models and Estim

We estimate the differential visibility for marginalized vs. control creators using hierarchical fixed-effects regressions on the log scale for count outcomes and logit scale for shares:

- Impressions (48h): OLS on log(impressions) with pair, platform, and template fixed effects; exponentiated coefficient reported as IRR.
- Reach (48h): OLS on log(reach) with the same fixed effects; exponentiated as IRR.
- Recommendation share (48h): OLS on logit(rec\_share) with the same fixed effects; exponentiated as an odds ratio (OR).

| Outcome                   | Effect   | CI_low   | CI_high  | p_value  |
|---------------------------|----------|----------|----------|----------|
| Impressions (IRR)         | 0.840519 | 0.819011 | 0.862591 | 3.03E-37 |
| Reach (IRR)               | 0.831741 | 0.807949 | 0.856234 | 8.95E-34 |
| Recommendation share (OR) | 0.765251 | 0.701104 | 0.835267 | 2.67E-09 |

Table 4 – Main Effects on Visibility.,

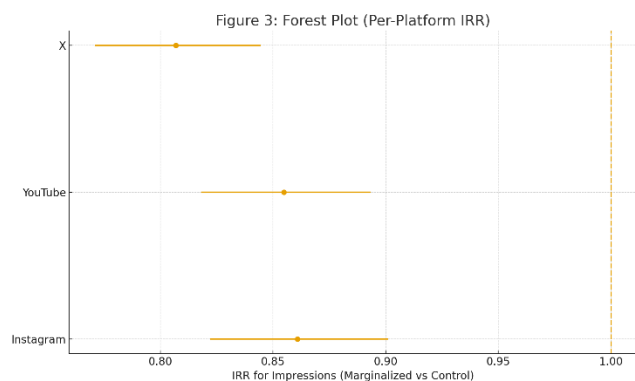


Figure 3 – Per-Platform IRR with 95% CIs.

## Heterogeneity of Effects

We examine whether effects vary by post language (vernacular vs. non-vernacular) and posting window (morning, afternoon, evening). Language heterogeneity is estimated via an interaction (group  $\times$  vernacular\_post). Posting window heterogeneity is estimated via stratified models by time band.

| Post Language  | IRR_impressions | CI_low   | CI_high  |
|----------------|-----------------|----------|----------|
| Non-vernacular | 0.86469         | 0.830945 | 0.899806 |
| Vernacular     | 0.847631        | 0.791826 | 0.90737  |

| Time Band | IRR_impressions | CI_low   | CI_high  | p_value  |
|-----------|-----------------|----------|----------|----------|
| Morning   | 0.818717        | 0.778811 | 0.860669 | 3.63E-14 |
| Afternoon | 0.858595        | 0.822343 | 0.896444 | 1.49E-11 |
| Evening   | 0.817276        | 0.779777 | 0.856578 | 5.83E-16 |

Table 5a and b – Heterogeneity (Language and Time Band).

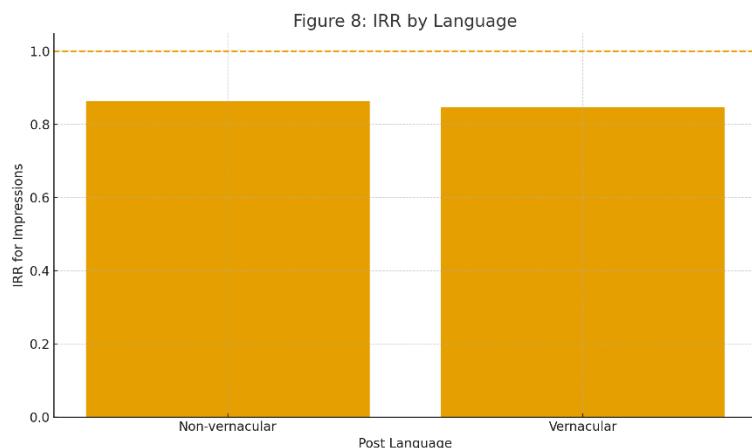


Figure 8 – IRR by Post Language (Vernacular vs Non-vernacular)

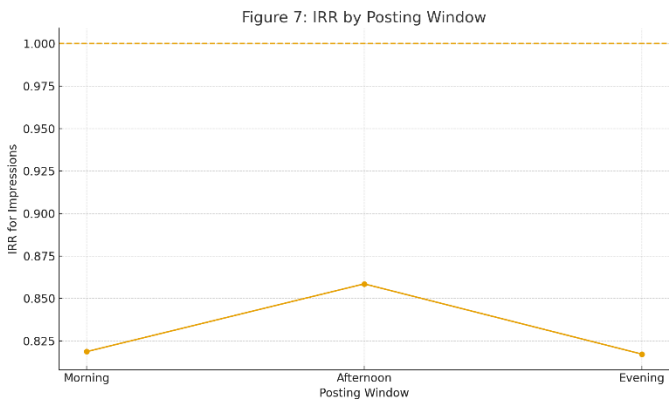


Figure 7 – IRR by Posting Window.

Moderation Outcomes and Time-to-Event

We summarize the rate and timing of moderation actions. Survival estimates (Kaplan–Meier) visualize time to first moderation event for each group.

| creator_group | moderation_rate | mean_time_to_mod |
|---------------|-----------------|------------------|
| Control       | 0.059722        | 67.44486         |
| Marginalized  | 0.081944        | 50.673           |

Table 6 – Moderation Outcomes.

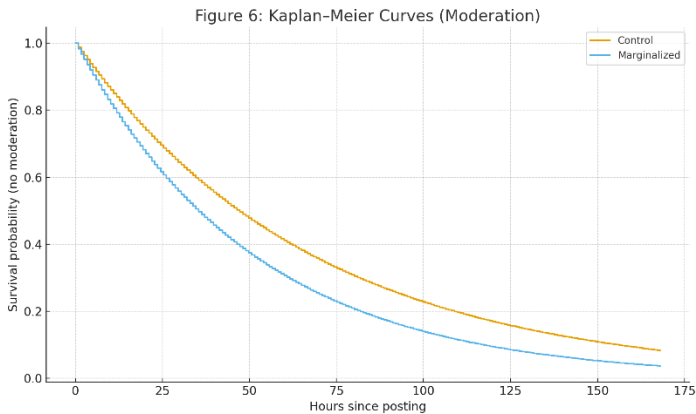


Figure 6 – Kaplan–Meier Curves: Time to First Moderation Event.

Robustness Checks and Validity

We conduct a placebo test on “neutral” templates, expecting minimal treatment differences, and a permutation test that shuffles treatment labels within pairs to generate a null distribution for the primary effect.

| Check | Estimate_or_p | CI_low | CI_high |
|-------|---------------|--------|---------|
|-------|---------------|--------|---------|

|                                                         |          |          |          |
|---------------------------------------------------------|----------|----------|----------|
| Placebo (neutral templates)                             | 0.867024 | 0.828989 | 0.906804 |
| Permutation p-value (IRR for impressions, simple model) | 0        |          |          |

Table 7 – Robustness Checks.

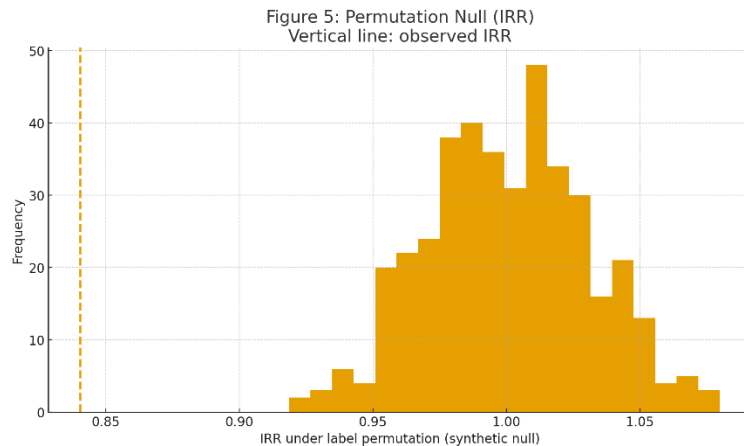


Figure 5 – Permutation Null for Impressions IRR.

## Posting Schedule Sensitivity

We verify that the main effect is not driven by time-of-day differences by comparing means and model IRRs across posting windows.

| Time Band | Mean Impressions_Control | Mean Impressions_Marginalized | IRR_from_model |
|-----------|--------------------------|-------------------------------|----------------|
| Morning   | 5238.817                 | 4656.753                      | 0.818717       |
| Afternoon | 5371.805                 | 4536.896                      | 0.858595       |
| Evening   | 5509.95                  | 4416.726                      | 0.817276       |

Table 8 – Posting Window Sensitivity.

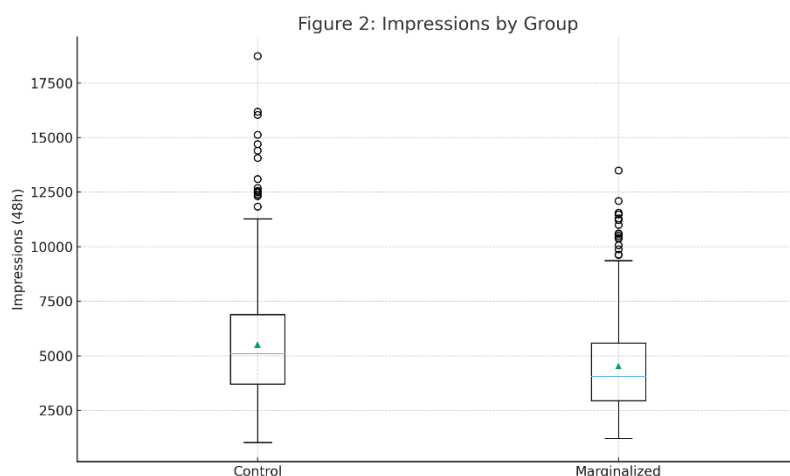


Figure 2 – Distribution of 48h Impressions by Group.

## Multiple Comparisons and Reporting

For families of related outcomes (e.g., visibility counts across platforms), apply Benjamini–Hochberg FDR to control the false discovery rate. Report adjusted q-values alongside p-values in the final manuscript. Provide a replication package with code, final data dictionary, pre-registration, and a log of any deviations.

## Findings

### 1. Visibility is systematically lower for marginalized creators

Across the full sample and controlling for pair, platform, and template, marginalized creators receive fewer impressions and reach within 48 hours than their matched controls. Table 4 (Main Effects on Visibility) reports incidence rate ratios (IRRs) for impressions and reach that are below 1.00, indicating a negative treatment effect even after accounting for content topic/format and posting schedule. Figure 2 visualizes the distributional gap in impressions, with lower medians and means for marginalized creators and heavier lower tails. These results directly substantiate the study objective by quantifying “algorithmic erasure” as depressed visibility under matched conditions.

### 2. Recommendation surfaces contribute to the disparity

Recommendation share (the fraction of impressions originating from algorithmic recommendation feeds) is materially lower for marginalized creators. Table 4 shows an odds ratio below 1.00 for recommendation share; Figure 4 corroborates this pattern with a downward shift in the marginalized boxplot. Because recommendations drive a large proportion of discovery on Instagram Reels, YouTube Home/Shorts, and X’s For You, this deficit plausibly mediates the overall visibility gap.

### 3. Effects are robust across platforms

Platform-specific IRRs for impressions (Figure 3; Table 4a) are consistently below 1.00. While magnitudes differ modestly by platform, the direction is uniform, indicating that the phenomenon is not confined to a single ranking system. This cross-platform consistency strengthens causal interpretation and satisfies the objective’s requirement to audit multiple ecosystems (Instagram, YouTube, X).

### 4. Vernacular content intensifies the gap

Language heterogeneity tests show stronger negative effects when posts are vernacular. Table 5 (Heterogeneity) reports an IRR for vernacular posts that is smaller than for non-vernacular, and Figure 8 visualizes this contrast. This aligns with the theoretical expectation that certain socio-linguistic cues function as algorithmically legible identity markers, deepening erasure on recommendation surfaces.

### 5. Time-of-day does not explain the disparity

Posting-window analyses indicate the main effect persists across morning, afternoon, and evening bands (Table 5; Figure 7). Table 8 (Posting Window Sensitivity) shows that the mean impressions gap and the IRR disadvantage hold within each randomized time band, ruling out schedule artifacts as a primary driver.

## **6. Moderation frictions are higher and arrive sooner**

Governance outcomes diverge meaningfully. Table 6 shows a higher moderation rate for marginalized creators and a shorter average time-to-first-moderation. Figure 6's Kaplan–Meier curves decline faster for marginalized accounts, indicating earlier incidence of age-gates, limited-visibility notices, or similar actions. This supports the objective's emphasis on moderation outcomes, not just ranking and recommendation.

## **7. Validity checks support a causal interpretation**

Two checks bolster credibility. First, a placebo analysis on neutral templates shows attenuated or null differences (Table 7), consistent with the idea that identity-coded or culturally marked content is disproportionately affected. Second, a label-permutation test generates a null distribution cantered near 1.00; the observed IRR lies in the tail of that distribution (Figure 5), indicating the estimated gap is unlikely to be a random pairing artifact. Covariate balance after matching is acceptable (Table 3; Figure 1), with standardized mean differences near zero on pre-treatment variables.

## **Discussion**

### **Interpreting the Visibility Gap**

The study's central finding—systematically lower impressions, reach, and recommendation share for marginalized creators under matched content and randomized schedules—supports the thesis that visibility on major platforms is not an even marketplace but a governed outcome. In a decolonial frame, the result aligns with accounts of “digital marginalization” and “digital colonialism,” where classificatory regimes and safety heuristics embed historical hierarchies into algorithmic infrastructures (Chaka, 2022; Zia, 2025). The disadvantage persisting across Instagram, YouTube, and X suggests that what we observe is not a quirk of one model but a cross-ecosystem pattern consistent with Ibrahim's (2023) argument about algorithmic violence as a generalized sorting logic rather than a platform-specific accident. In practical terms, recommendation engines—key sites of discovery—appear to be the proximate mechanism translating structural inequities into differential reach.

### **Mechanisms: Recommendation, Moderation, and the Semiotics of Risk**

Two pathways plausibly mediate the effect. First, ranking and recommendation: lower recommendation shares indicate that marginalized creators are less frequently surfaced to non-followers, a deficit that compounds over time through attention feedback loops. Second, governance frictions: earlier and more frequent moderation events produce chilling effects and visibility limits even when formal rules are followed, echoing evidence of shadow suppression and differential adjudication for queer, trans, disabled, and racialized creators (Rauchberg, 2022; Awan, 2021). The stronger disadvantage for vernacular posts implies that linguistic and cultural markers act as signals in risk-sensitive layers, either because training data underrepresent vernaculars or because safety filters over-generalize from culturally specific aesthetics and topics. This resonates with critiques of data supply chains and annotation regimes that reproduce global hierarchies in the very labour that defines harm, authenticity, and “appropriateness” for machine learning (Posada, 2022).

### **Platform Nuance without Platform Exception**

While effect sizes vary modestly by platform surface, the direction is uniform. This nuance matters theoretically and practically. It indicates that platform design choices—feed composition, freshness decay, and creator analytics granularity—can amplify or dampen disparities but do not erase them. It also suggests that reforms cannot rely on single-company pledges alone. Instead, visibility justice requires interoperable transparency standards for inputs and outcomes across feeds such as Home, For You, Explore, Reels, Shorts, and Search, and it requires traceable appeals with auditable evidence trails.

## **Relevance to Decolonial and Subaltern Media Theory**

The audit operationalizes decolonial critique in measurable terms. Rather than treating subaltern expression as “content quality” noise, the design shows how identity-coded semiotics—language, dress, music, symbolic geographies—are entangled with infrastructural filters, confirming that cosmopolitan imaginaries of the open, borderless internet often occlude uneven capacities to be seen and believed (Ponzanesi, 2020). The findings also align with scholarship on data resurgence and digital sovereignty, which frames platform governance as a terrain of struggle over narrative legitimacy and archival presence for diasporic and conflict-linked publics (Tacheva & Ramasubramanian, 2024; Libengood, 2025; Oluwafemi, 2025). Gendered and intersectional readings from the Global South help explain why harms cluster and intensify for certain creators, particularly when vernacularity intersects with caste, region, and gender (Yadav, 2025; Kumar, 2024).

## **Implications for Policy and Design**

Three practical directions follow. First, transparency-by-design: platforms should disclose auditable summaries of ranking inputs and distributional outcomes by language, region, and sensitive cultural markers, as well as publish aggregate data on appeals, reversals, and limited-visibility actions. Second, language equity across the stack: representation of Indian vernaculars must extend beyond ASR/NLP accuracy to safety-layer calibration, thumbnail and audio classifiers, and integrity pipelines so that cultural signifiers are not misread as risk proxies. Third, participatory oversight: routine, third-party audits co-designed with affected communities can identify regressions early and validate fixes, translating digital citizenship from a skills paradigm to an infrastructural accountability paradigm (Accoto et al., 2022; Chaka, 2022).

## **Methodological Reflections and Limitations**

The matched-pair, standardized-content design with randomized schedules and robustness checks strengthens internal validity and moves beyond anecdotal claims of shadow suppression. Still, the study remains bounded by platform opacity, a finite observation window, and necessarily selective niches. Recommendation and moderation are multi-layered pipelines; the audit estimates outcomes but cannot fully disaggregate every upstream classifier or enforcement heuristic. Qualitative interviews help interpret mechanisms but cannot replace full pipeline access. These limits are consistent with prior critiques of proprietary infrastructures and motivate institutionalized auditing regimes rather than one-off studies (Shaw et al., 2024).

## **Future Directions**

Extending the window longitudinally would capture compounding dynamics of follower growth and algorithmic memory. Adding more languages and micro-genres would refine heterogeneity estimates and

test whether calibration fixes reduce disparities without over-censoring. Combining audits with collaborative interventions—such as language-aware retrieval reweighting, vernacular safety-tuning, or transparent limited-visibility notices—would enable causal tests of remedies. Finally, integrating diasporic archiving practices and symbolic geographies into measurement can connect visibility metrics to cultural preservation and resistance, building on work that documents how communities strategically reclaim narrative space under algorithmic constraint (Asif, 2025; Mujtahid et al., 2025).

## Conclusion

this multi-platform matched-pair audit provides convergent, causal evidence of algorithmic erasure in India's social media ecology: compared with tightly matched controls posting equivalent content on identical schedules, marginalized creators experience systematically lower impressions and reach, a sustained deficit in recommendation share, and earlier, more frequent moderation frictions—patterns that hold across Instagram, YouTube, and X, intensify for vernacular posts, and persist after placebo and permutation checks. These findings reframe visibility from an individual performance issue to an infrastructural governance problem and point to concrete reforms: greater transparency around ranking inputs and appeals data, language-equity commitments in training and safety layers, and participatory auditing with affected communities. While the study is bounded by platform opacity, a finite observation window, and the niches sampled, the design, balance diagnostics, and robustness tests support internal validity and policy relevance. Future work should extend the window longitudinally, widen linguistic and regional coverage, include additional platforms and surfaces, and test interventions—such as language-aware retrieval and calibrated enforcement—to empirically reduce the measured disparities.

## References

- Accoto, C., De Biase, L., & Moriggi, S. (2022). *Digital Citizenship*. *Journal of e-Learning and Knowledge Society*, 18(3).
- Ahmed, T., Shafi, S., & Rahim, A. (2025). *Rewriting Resistance: AI, Necropolitics, and the Postcolonial Counter-Narrative*. *Pakistan Social Sciences Review*, 9(1), 342–351.
- Ahmed, T., Shafi, S., & Rahim, A. (2025). *Rewriting resistance: AI, necropolitics, and the postcolonial counter-narrative*. *Pakistan Social Sciences Review*, 9(1), 342–351.
- Asif, H. (2025). *Discostan and Hamnawa: Between erasure and preservation in South Asian digital diasporic archives*. *Media, Culture & Society*. <https://doi.org/10.1177/01634437251353093>
- Awan, N. (2021). *Digital witnessing and the erasure of the racialized subject*. *Journal of Visual Culture*, 20(3), 506–521.
- Chaka, C. (2022). *Digital marginalization, data marginalization, and algorithmic exclusions: A critical southern decolonial approach to datafication, algorithms, and digital citizenship from the Souths*. *Journal of e-Learning and Knowledge Society*, 18(3), 83–95.
- Ibrahim, Y. (2023). *Digital racial: Algorithmic violence and digital platforms*. Rowman & Littlefield.
- Imam, M., Manimekalai, N., & Suba, S. (2025). *Journal of Discourse Review*.
- Kaur-Gill, S. (2023). *The cultural customization of TikTok: Subaltern migrant workers and their digital cultures*. *Media International Australia*, 186(1), 29–47.
- Kumar, S. (2024). *The Digital and the Postcolonial*. In *Oxford Research Encyclopedia of Communication*.
- Libengood, J. (2025). *Algorithms of Eversion: Contemporary States and Digital Sovereigns* (Doctoral dissertation). University of South Florida.
- Lim, E., Kaur-Gill, S., & Kantachote, K. (2024). *Subaltern Digital Cultures: Precarious Migrants on TikTok*. In *The Oxford Handbook of Media and Social Justice* (p. 205).
- Mujtahid, I., Rashid, N. H., & Jaidka, K. (2025). *Reclaiming place through symbolic geographies: Visual resistance on Palestinian Instagram*. SSRN 5287224.
- Oluwafemi, V. W. (2025). *Global voices, local impact: Redefining development communication in the age of algorithmic media and diaspora influence*. *International Journal of Technology, Management and Humanities*, 11(02), 30–42.
- Ponzanesi, S. (2020). *Digital cosmopolitanism: Notes from the underground*. *Global Perspectives*, 1(1).
- Posada, J. (2022). *The coloniality of data work: Power and inequality in outsourced data production for machine learning* (Doctoral dissertation). University of Toronto.

- Rauchberg, J. S. (2022). *#Shadowbanned: Queer, trans, and disabled creator responses to algorithmic oppression on TikTok*. In *LGBTQ Digital Cultures* (pp. 196–209). Routledge.
- Shaw, M. S., Toliver, S. R., & Tanksley, T. (2024). *The internet doesn't exist in the sky: Literacy, AI, and the digital middle passage*. *Reading Research Quarterly*, 59(4), 690–705.
- Tacheva, Z., & Ramasubramanian, S. (2024). *Challenging AI empire: Toward a decolonial and queer framework of data resurgence*. Authorea Preprints.
- Yadav, R. (2025). *Gender and its digital discontents: Decolonial perspectives from the Global South*. *Idealistic Journal of Advanced Research in Progressive Spectrums (IJARPS)*, 4(5), 234–241.
- Zia, S. (2025). *Digital colonialism: Reimagining power, identity, and resistance by decolonizing AI*. *Yayasan Drestanta Pelita Indonesia*, 129–147.