

Eduard-Claudiu GROSS*
Anda-Ștefana DOBRE**

Stitching the Truth: Investigating TikTok's Stitch Feature as a Native Debunking Practice in the Romanian Digital Sphere

Abstract

This study examines stitch-based debunking as a platform-native practice of misinformation correction on Romanian TikTok. As short-form video platforms become increasingly central to information consumption, users are appropriating native technical features, particularly TikTok's stitch function, to challenge and reframe misleading content. Drawing on a mixed-methods analysis of stitch-based debunking videos, the study combines qualitative content analysis, engagement metrics, and computational text analysis to examine how corrective practices are performed, circulated, and received. The findings show that stitch-based debunking operates primarily as a peer-to-peer corrective mechanism embedded within platform discourse rather than as an extension of institutional fact-checking. Debunking content frequently targets political, health-related, and gendered misinformation and relies heavily on moral framing and direct address. While informative corrections are most common, content employing humor or sarcasm tends to attract higher levels of engagement, highlighting a tension between informational clarity and platform visibility. Audience responses are generally neutral to supportive, and extended stitch chains are rare, suggesting that algorithmic distribution plays a greater role in amplification than collaborative chaining. Overall, the findings suggest that grassroots debunking may constitute a meaningful component of contemporary information ecosystems and point to the importance of platform affordances in shaping participatory forms of civic media literacy in non-Anglophone contexts.

Keywords: TikTok, stitch feature, misinformation, debunking, platform affordances, participatory media literacy

1 Introduction

The proliferation of misinformation across digital platforms has emerged as one of the defining challenges of contemporary democratic societies (Benkler et al., 2018; Iyengar & Hahn, 2009; Lelkes, 2016; Sunstein, 2017). While a substantial body of scholarship has ex-

* Lucian Blaga University of Sibiu, eduard.gross@ulbsibiu.ro,
ORCID: <https://orcid.org/0000-0002-6842-1173>, corresponding author

** Lucian Blaga University of Sibiu, andastefana.dobre@ulbsibiu.ro,
ORCID: <https://orcid.org/0009-0005-5550-8439>

amined the production, dissemination, and psychological effects of false information (Corbu et al., 2020; Hamелеers et al., 2018; Pennycook & Rand, 2019; Thorson, 2016; Vosoughi et al., 2018), comparatively less attention has been paid to the corrective practices that arise organically within platform communities themselves (Marwick et al., 2025; Papacharissi, 2015). This study addresses that gap by examining how users on TikTok, a platform now central to youth media consumption globally (Cușnir, 2025), mobilize native platform affordances to respond to and contest misinformation. Specifically, the analysis focuses on TikTok's "stitch" feature, which allows users to incorporate up to five seconds of another user's video into their own content, producing a direct visual and rhetorical juxtaposition between an original claim and a response (Zulli & Zulli, 2022). This affordance enables a form of dialogic correction that differs substantially from institutional fact-checking practices (Bucher & Helmond, 2018; Gibson, 2015; Gillespie, 2018). Rather than operating through separate expert-driven channels, stitch-based debunking embeds corrective discourse directly within the same viewing experience as the misinformation itself (Graves, 2016), positioning ordinary users as active agents in information verification.

Beyond the mere presence of correction, stitch-based debunking varies widely in form and style. Users draw on diverse rhetorical, visual, and tonal strategies, ranging from humor and irony to moral appeals and evidentiary rebuttals, consistent with prior work on participatory media and memetic communication (Gray et al., 2009; Shifman, 2014; Stancea et al., 2025). Because stitches can generate extended chains of responses rather than isolated counterclaims, they also create evolving discursive spaces in which meanings, arguments, and frames are negotiated over time (Gamson & Modigliani, 1989). Stitch-based corrections invite immediate and visible feedback through comments, likes, and shares, while algorithmic distribution mechanisms shape the reach and visibility of corrective content (Aral, 2021; Berger, 2016). The Romanian digital context offers a particularly salient site for investigating these dynamics. Romania ranks among the European Union member states with the highest levels of concern regarding misinformation, with surveys indicating that more than 63% of Romanians encounter content they perceive as false or misleading on a weekly basis (Dumitrescu, 2023). At the same time, institutional fact-checking infrastructures remain comparatively underdeveloped, creating conditions in which grassroots corrective practices may play an outsized role. The Romanian TikTok ecosystem thus provides a valuable opportunity to examine how platform-native debunking emerges and circulates in contexts where formal verification mechanisms are limited. By analyzing stitch-based misinformation correction on TikTok, this study contributes to emerging scholarship at the intersection of platform studies, participatory culture, and misinformation research. Through attention to communicative strategies, discursive evolution, audience reception, and patterns of spread, the analysis aims to shed light on how technical affordances shape civic practices of information verification. More broadly, the findings carry implications for media literacy education, platform governance, and our understanding of how democratic deliberation adapts within participatory media environments.

2 Literature review

2.1 Misinformation in the platform era

The study of misinformation has undergone substantial transformation in the platform era, shifting from concerns about traditional media bias toward analysis of networked information disorder. Wardle and Derakhshan's (2017) influential framework distinguishes between misinformation (false content shared without harmful intent), disinformation (deliberately fabricated or manipulated content), and malinformation (genuine information shared to cause harm). This typology has proven particularly useful for understanding the diverse ecosystem of problematic content circulating on social media platforms, where intentionality is often difficult to determine and content frequently mutates through sharing and remixing. Research on the psychological dimensions of misinformation has identified several cognitive mechanisms that contribute to its persistence and influence. The continued influence effect describes the tendency for misinformation to shape reasoning even after correction, particularly when the false information aligns with prior beliefs or fills explanatory gaps in understanding (Lewandowsky et al., 2012). Related work on motivated reasoning suggests that corrections may prove ineffective, or even counterproductive, when they challenge identity-relevant beliefs, potentially triggering „backfire effects“ that strengthen rather than weaken false convictions (Nyhan & Reifler, 2010). However, more recent meta-analytic work has questioned the prevalence of backfire effects, suggesting that corrections generally reduce misperceptions even when they do not entirely eliminate them (Wood & Porter, 2019). The platform-specific dynamics of misinformation have received increasing scholarly attention. Research has documented how algorithmic recommendation systems can create „filter bubbles“ and „echo chambers“ that reinforce existing beliefs while limiting exposure to corrective information (Pariser, 2012). However, empirical investigations have produced mixed evidence regarding the prevalence and effects of such phenomena, with some studies suggesting that algorithmic exposure to cross-cutting content may actually exceed offline networks (Bakshy et al., 2015). The debate highlights the importance of examining misinformation dynamics within specific platform contexts rather than assuming universal patterns across digital environments.

2.2 Platform affordances and TikTok

The concept of affordances, originally developed in ecological psychology, has been productively adapted to analyze how technical features of digital platforms enable, constrain, and shape social action. Bucher and Helmond (2018) distinguish between low-level technical affordances (features like commenting, sharing, or stitching) and high-level communicative affordances (the social practices these features enable). This framework proves particularly useful for understanding how platform-specific tools become appropriated for purposes not necessarily anticipated by their designers. TikTok presents a distinctive configuration of affordances that differentiate it from earlier social media platforms. The platform's algorithmic „For You“ feed prioritizes content-based recommendations over social graph connections, enabling rapid viral distribution for creators without established followings (Zulli & Zulli, 2022). The emphasis on audiovisual content, combined with features like duetting and stitching, creates affordances for dialogic engagement that differ substantially from text-based platforms.

These features have been theorized as enabling „networked“ or „distributed“ forms of content creation in which meaning emerges through chains of response and recontextualiza-

tion (Abidin, 2020). The stitch feature specifically allows users to incorporate up to five seconds of another user's video at the beginning of their own content. Unlike duets, which display both videos simultaneously in split-screen format, stitches create a sequential juxtaposition that positions the original content as a prompt for response. This temporal structure proves particularly suited to corrective interventions, as it mirrors the claim-then-correction format recommended in debunking scholarship (Lewandowsky et al., 2020). The five-second limitation also forces creators to distill misinformation claims to their essential elements, potentially reducing the risk of inadvertent amplification through extended repetition.

2.3 Participatory culture and civic media literacies

The emergence of user-generated debunking can be situated within broader scholarly conversations about participatory culture and civic media literacies. Jenkins and colleagues' foundational work on participatory culture emphasized how digital platforms enable forms of creative production and circulation previously reserved for professional media industries (Jenkins, 2006; Jenkins et al., 2016). While initial scholarship emphasized entertainment-oriented practices like fan fiction and remix culture, subsequent work has examined how participatory affordances are appropriated for explicitly civic purposes. Mihailidis (2018) introduces the concept of „civic media literacies“ to describe the competencies required for effective participation in digitally mediated public life. Beyond traditional media literacy skills of critical analysis, civic media literacies encompass capacities for production, circulation, and collaborative meaning-making. This framework suggests that debunking practices represent not merely consumption of verified information, but active participation in collective processes of knowledge construction and epistemic maintenance. Related scholarship on „produsage“ (Bruns, 2009a) and „gatewatching“ (Bruns, 2009b) has examined how the boundaries between producers and consumers of information have blurred in digital environments. Gatewatching describes practices through which users monitor, select, and republish content from diverse sources, functioning as distributed curators rather than traditional gatekeepers (Bruns, 2018). Stitch-based debunking can be understood as a form of critical gatewatching in which users not only select and circulate content but explicitly contest and reframe it.

2.4 Counter-speech and networked debunking

Research on counter-speech, defined as responses to harmful speech that seek to dilute, challenge, or undermine its message (Benesch, 2016), provides additional theoretical resources for understanding stitch-based debunking. Benesch distinguishes between organized counter-speech campaigns and organic responses emerging from affected communities, noting that the latter may prove more authentic and persuasive precisely because they lack institutional sponsorship (Benesch, 2016). This distinction maps productively onto the difference between professional fact-checking organizations and grassroots debunking practices on platforms like TikTok.

The concept of „networked debunking“ describes how corrective information circulates through social networks, potentially reaching audiences that institutional fact-checks might not access (Margolin et al., 2018). Research in this area has examined how social endorsement affects the credibility and reach of corrective information, finding that corrections shared by known contacts may prove more persuasive than those from unfamiliar expert sources. The

peer-to-peer character of stitch-based debunking – in which ordinary users correct content produced by other ordinary users – may thus confer advantages in terms of perceived authenticity and social proof. However, scholarship has also identified challenges associated with decentralized correction (Tandoc et al., 2020; Vraga & Bode, 2017). Without editorial standards or verification procedures, user-generated debunking may itself contain inaccuracies or introduce new forms of misinformation. The absence of institutional accountability raises questions about quality control, while the potential for debunking to become performative – oriented toward engagement and self-presentation rather than epistemic goals – suggests that not all corrective content serves genuinely informational purposes (Tandoc et al., 2020).

2.5 Debunking effectiveness and best practices

A substantial body of experimental research has examined the effectiveness of different debunking strategies, generating evidence-based recommendations for correction design. The „truth sandwich“ approach, popularized by linguist George Lakoff, recommends leading with accurate information, briefly acknowledging the false claim, and then repeating the correction, thereby minimizing repetition of misinformation while maximizing exposure to accurate information (Lakoff et al., 2004). Related work suggests that explanatory corrections, which provide alternative accounts rather than merely labeling claims as false, prove more effective at reducing continued influence effects (Lewandowsky et al., 2020).

Research on source credibility has produced nuanced findings regarding the optimal characteristics of correction sources. While expert sources generally enhance correction effectiveness, perceived bias or political motivation can undermine credibility even among ostensibly expert correctors (Vraga & Bode, 2017). Peer corrections – from sources similar to the audience rather than distant authorities – may prove effective precisely because they avoid triggering motivated resistance associated with out-group messengers. These findings suggest that the peer-to-peer character of TikTok debunking may constitute an advantage rather than a limitation. The role of emotional and rhetorical framing in correction effectiveness remains contested. Some research suggests that humor and ridicule can enhance engagement with corrective content and reduce defensive reactions to challenging information (Brewer & McKnight, 2017). However, other work cautions that humor may trivialize serious misinformation or alienate audiences who feel mocked rather than informed. The tension between engagement optimization and informational effectiveness represents a key challenge for understanding platform-native debunking practices.

2.6 The Romanian Context

Romania presents a distinctive context for examining misinformation dynamics and corrective responses. As a post-communist society that underwent rapid democratization and digital transformation, Romania exhibits particular patterns of institutional trust and media consumption. Surveys consistently indicate lower levels of trust in traditional media compared to Western European averages, combined with high rates of social media adoption particularly among younger demographics (Newman et al., 2025). The Romanian misinformation landscape has been shaped by both domestic political dynamics and transnational information flows. Despite these challenges, Romania's fact-checking infrastructure remains relatively underdeveloped. While organizations like Factual.ro and initiatives from mainstream

media outlets provide institutional verification services, their reach and visibility remain limited compared to the scale of misinformation circulation. This gap creates space for alternative forms of correction to emerge, including the platform-native debunking practices examined in this study.

2.7 Research Questions

Building on the theoretical foundations outlined above, this study addresses five interrelated research questions. First, given the distinctive affordance structure of TikTok's stitch feature and its potential for dialogic correction, the study asks: RQ1: How is the TikTok stitch function used to respond to and correct misinformation? Second, because debunking content varies in its rhetorical, visual, and tonal composition, and these variations may shape platform visibility: RQ2: Which rhetorical, visual, and tonal strategies are associated with higher engagement? Third, to examine whether stitch chains produce evolving argumentative trajectories: RQ3: How does the discourse evolve within stitch chains? Fourth, to assess how audiences receive peer-led correction: RQ4: How do TikTok audiences respond to stitch-based debunks? Finally, to understand the spread dynamics of corrective content: RQ5: To what extent do stitch chains contribute to viral spread?

3 Methodology

3.1 Research Design

This study adopts a mixed-methods research design combining quantitative analysis of engagement patterns with qualitative content analysis and computational text analysis. The approach is informed by an abductive analytical logic, understood as iterative movement between empirical observations and theoretical frameworks that enables the development of novel conceptual insights (Tavory & Timmermans, 2014). Rather than testing pre-specified hypotheses, the study aims to develop grounded understanding of how stitch-based debunking operates as a situated practice within the TikTok ecosystem. TikTok is approached not merely as a repository of content but as a socio-technical system whose affordances shape visibility, interaction, and meaning-making. The study therefore treats debunking as a performative and participatory practice embedded in platform-specific norms and technical features. This orientation aligns with platform studies scholarship that emphasizes the co-constitutive relationship between technical infrastructures and social practices (Pilipets & Chao, 2025).

3.2 Data sampling and collection

3.2.1 Sampling Strategy

Data were collected using a purposive, non-probabilistic sampling strategy appropriate for exploratory research on emergent digital practices. Given the absence of a comprehensive archive or reliable search mechanism for stitch-based debunking content on TikTok, sampling relied on iterative platform exploration combined with snowballing techniques. Initial cases were identified through: (1) keyword and hashtag searches related to misinformation, fact-checking, and debunking in Romanian; (2) algorithmic recommendations within Tik-

Tok's „For You“ feed; and (3) visible stitch chains linking misleading content with corrective responses. Subsequent cases were selected by tracing stitch chains and recurring creator networks, allowing the corpus to expand based on theoretical relevance and variation rather than numerical representativeness. This approach was designed to capture diversity in debunking practices, rhetorical styles, and creator positioning.

3.2.2 Inclusion Criteria

Videos were included in the corpus if they met the following criteria: (1) the video explicitly employed TikTok's stitch feature; (2) the stitched segment referenced content that was misleading, false, or framed as such by the creator; (3) the creator articulated an explicit intent to correct, contextualize, or challenge the original claim; (4) the video was produced by a Romanian-based creator or addressed issues salient to the Romanian informational environment; and (5) engagement metrics were publicly visible at the time of collection. Videos that merely reacted to content without discernible corrective intent were excluded.

3.2.3 Final Corpus

The final corpus consists of 80 TikTok stitch videos produced by 43 unique creators, collected between April and September 2025. For each video, the following data were recorded at the time of collection: metadata (publication date, duration), creator profile information (username, follower count, bio content), engagement metrics (views, likes, comments, shares, saves), captions, hashtags, and sound usage. Additionally, the top five publicly visible user comments were collected for each video to enable qualitative analysis of audience reception. Full transcripts were produced for 70 videos with sufficient audio clarity for transcription.

3.3 Analytical Framework

Data analysis proceeded through multiple iterative rounds combining close viewing and reading with systematic coding. The analytical framework is structured around four interrelated dimensions, each anchored in established bodies of literature:

3.3.1 Platform Affordances and Multimodal Practices

Drawing on theories of social media affordances (Bucher & Helmond, 2018; Davis & Chouinard, 2016), this dimension examines how TikTok's native features, including the stitch function, audiovisual layering, sound reuse, text overlays, and visual performance, structure debunking practices. These affordances are treated as enabling and constraining conditions that shape how corrective messages are articulated, stylized, and made visible.

3.4 Debunking Strategies and Counter-Speech Forms

Informed by research on misinformation correction and counter-speech (Lewandowsky et al., 2017; Wardle & Derakhshan, 2017), this dimension focuses on rhetorical strategies and tonal framings used to challenge misleading content. Coding captures variations including informational fact-checking, satire and ridicule, moral framing, appeals to authority, and experiential or anecdotal correction. These categories were derived deductively from established typologies in misinformation correction research (Lewandowsky et al., 2020; Vraga et al.,

2019) and refined inductively through iterative pilot coding. Informational fact-checking refers to responses that present verifiable evidence or data to contradict false claims. Satire and ridicule involve the use of humor, irony, or mockery to undermine the credibility of misleading content. Moral framing positions correction within ethical or normative arguments appealing to shared values. Appeals to authority invoke professional expertise, institutional credibility, or credentials. Experiential or anecdotal correction draws on personal experience or testimony as counter-evidence. These strategies are analyzed as distinct communicative logics rather than evaluated for effectiveness. A detailed operationalization of all coding categories is provided in Appendix A.

3.4.1 Creator Positioning and Participatory Agency

Drawing on scholarship on participatory culture, presumption, and civic media literacies (Jenkins et al., 2016; Mihailidis, 2018), this dimension examines how creators position themselves as civic actors, performers, or hybrid agents. Attention is paid to motivational cues, identity performance, and repeated engagement with debunking practices as forms of peer-driven civic participation.

3.4.2 Audience Response and Reception

This dimension analyzes patterns of engagement and sentiment in audience responses, examining how comment sections function as sites of deliberation, affirmation, or contestation. Coding of comment sentiment enables assessment of whether debunking content provokes supportive, hostile, or mixed reactions.

3.5 Computational Text Analysis

To complement qualitative analysis, computational text analysis was applied to the corpus of transcribed video content. Latent Dirichlet Allocation (LDA) topic modeling was employed to identify latent thematic structures in debunking discourse. LDA is a generative probabilistic model that represents documents as mixtures of topics, where each topic is characterized by a distribution over words (Blei et al., 2003). The model was implemented using scikit-learn in Python, with preprocessing including stopword removal (for both Romanian and English), lowercasing, and removal of non-alphabetic characters. Model parameters were selected based on coherence scores and interpretability of resulting topics. The final five-topic model was validated through examination of topic keywords and manual review of high-loading documents for each topic. t-SNE (t-distributed Stochastic Neighbor Embedding) dimensionality reduction was applied to TF-IDF vectorized transcripts to visualize document clustering in two-dimensional space.

3.6 Statistical Analysis

Quantitative analysis of engagement patterns employed non-parametric statistical tests appropriate for the non-normal distribution's characteristic of social media metrics. Kruskal-Wallis H tests were used to assess differences in engagement across categorical variables (e.g., tone, topic), while Spearman's rank correlation coefficients measured associations be-

tween continuous variables. All statistical analyses were conducted in Python using SciPy and statsmodels libraries, with the significance threshold set at $\alpha = .05$.

3.7 Ethical Considerations

All analyzed content was publicly available at the time of collection. No direct interaction with creators or commenters occurred during the research process. In reporting findings, usernames are referenced only when necessary for analytical purposes, and the study avoids reproducing content that could cause reputational harm to individual creators. The purposive and snowball-based sampling strategy does not permit statistical generalization to the broader population of TikTok content, a limitation acknowledged in the discussion of findings.

3.8 Codebook Development

A comprehensive codebook was developed to guide systematic content analysis of the video corpus. The codebook operationalizes 37 variables across multiple dimensions including: video metadata (duration, publication date), creator characteristics (follower count, profile type, political positioning), content features (misinformation topic, origin of original claim, stitch chain position), rhetorical elements (tone, strategies, visual techniques), and audience response (comment sentiment, presence of debate). The codebook was refined through pilot coding of a subsample of videos, with ambiguous categories clarified through discussion and iterative revision. Inter-coder reliability was assessed on a subset of 15 videos, achieving acceptable agreement levels across primary coding categories.

4 Results

This section presents the findings from the analysis of 80 TikTok stitch-based debunking videos produced by 43 unique creators in the Romanian digital ecosystem. The results are organized according to the five research questions guiding this study, integrating quantitative engagement metrics, qualitative content coding, and computational text analysis including Latent Dirichlet Allocation (LDA) topic modeling. All videos were collected between April and September 2025, with a combined reach of over 5.6 million views.

4.1 Sample Characteristics

The analytical corpus comprises 80 TikTok videos explicitly employing the stitch feature for misinformation correction. Table 1 presents the descriptive statistics for the sample.

Table 1. Sample Descriptive Statistics

Variable	Value
Total videos analyzed	80
Unique creators	43
Collection period	April – September 2025
Total views	5,640,788
Mean views per video	72,318 (SD = 105,254)
Median views	22,850
Mean video duration	117.9 seconds (SD = 80.7)
Median duration	97.5 seconds

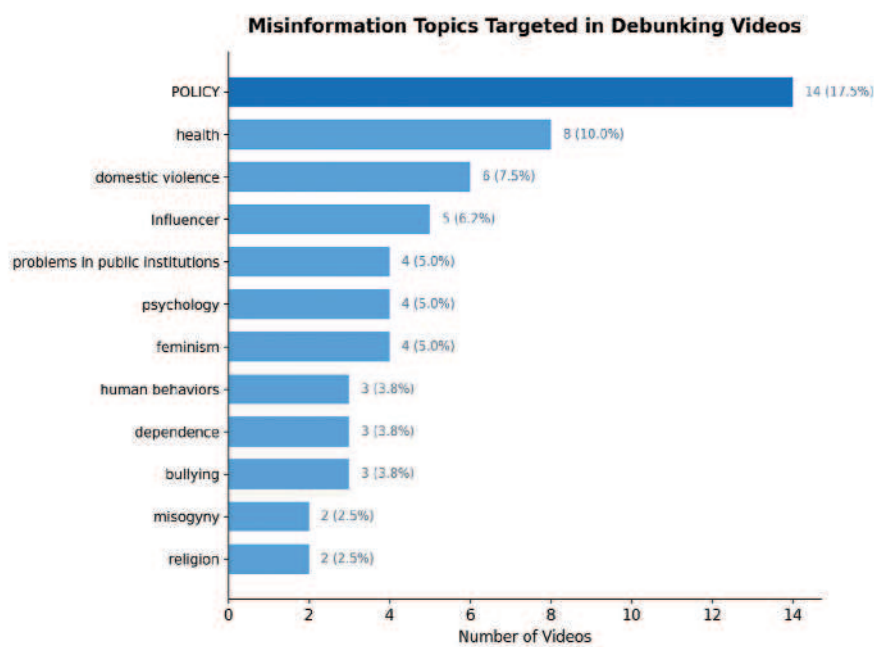
The substantial gap between mean and median views (72,318 vs. 22,850) indicates a right-skewed distribution characteristic of social media engagement, where a minority of videos achieve viral reach while the majority receive moderate attention.

4.2 RQ1: How is the TikTok Stitch Function Used to Respond to and Correct Misinformation?

4.2.1 Misinformation Topics Targeted

Analysis of the thematic content revealed that debunking efforts addressed a diverse range of misinformation topics, with political content emerging as the most frequently targeted category (Figure 1).

Figure 1. Misinformation topics targeted in debunking videos

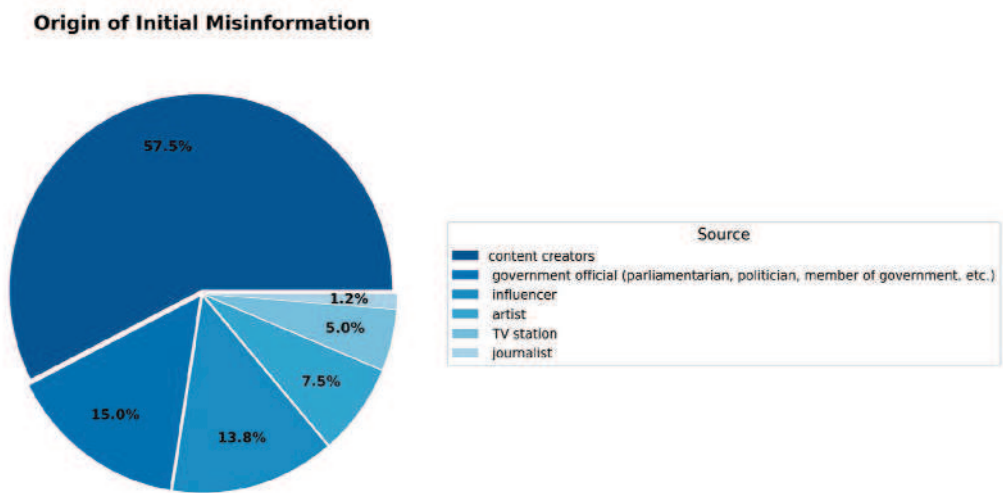


Political misinformation constituted the largest single category (17.5%), encompassing content related to electoral processes, political figures, and policy debates. Health-related misinformation ranked second (10.0%), followed by content addressing domestic violence narratives (7.5%). The prevalence of gender-related topics – including domestic violence, feminism, and misogyny – collectively represented a substantial portion of debunking activity, reflecting the salience of gender discourse in the Romanian digital public sphere.

4.2.2 Origins of Misinformation

Examination of the sources from which misinformation originated revealed that peer-produced content constitutes the primary target of stitch-based correction efforts (Figure 2).

Figure 2. Origin of initial misinformation sources

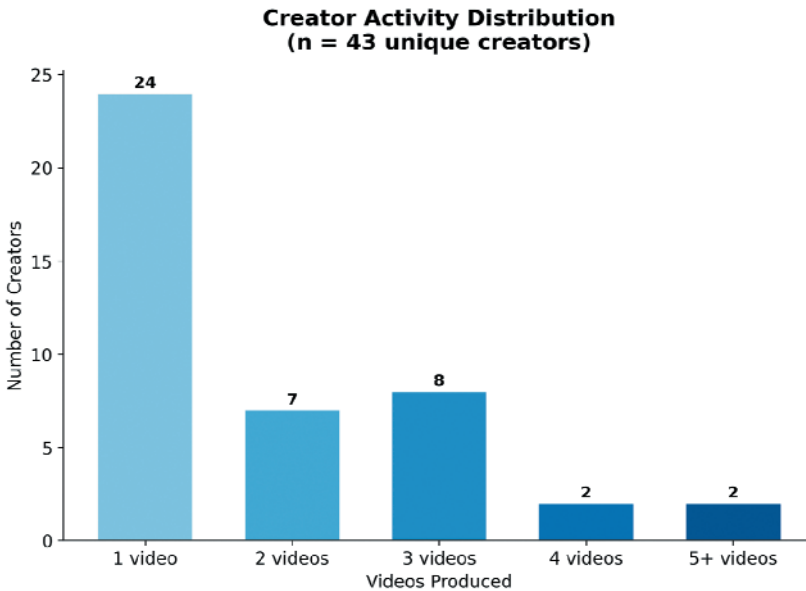


The predominance of content creators as misinformation sources (57.5%) suggests that stitch-based debunking operates primarily as a peer-to-peer correction mechanism within the platform ecosystem, rather than as a response to institutional media or official sources. Content creators are defined here as users who produce original content without a verified public profile or institutional affiliation, as distinct from influencers who maintain established followings and monetized presences. Government officials and politicians represented the second most common source (15.0%), indicating that political speech constitutes a significant target for grassroots fact-checking.

4.2.3 Creator Activity Patterns

Analysis of creator participation revealed varying levels of engagement with debunking activity (Figure 3).

Figure 3. Distribution of creator activity (videos per creator)



Most creators ($n = 25$, 58.1%) produced only a single debunking video in the corpus, while a smaller group of dedicated debunkers contributed multiple videos. This pattern suggests that stitch-based debunking attracts both occasional participants responding to specific misinformation and more systematic correctors who engage regularly with platform discourse.

4.2.4 Stitch Chain Structure

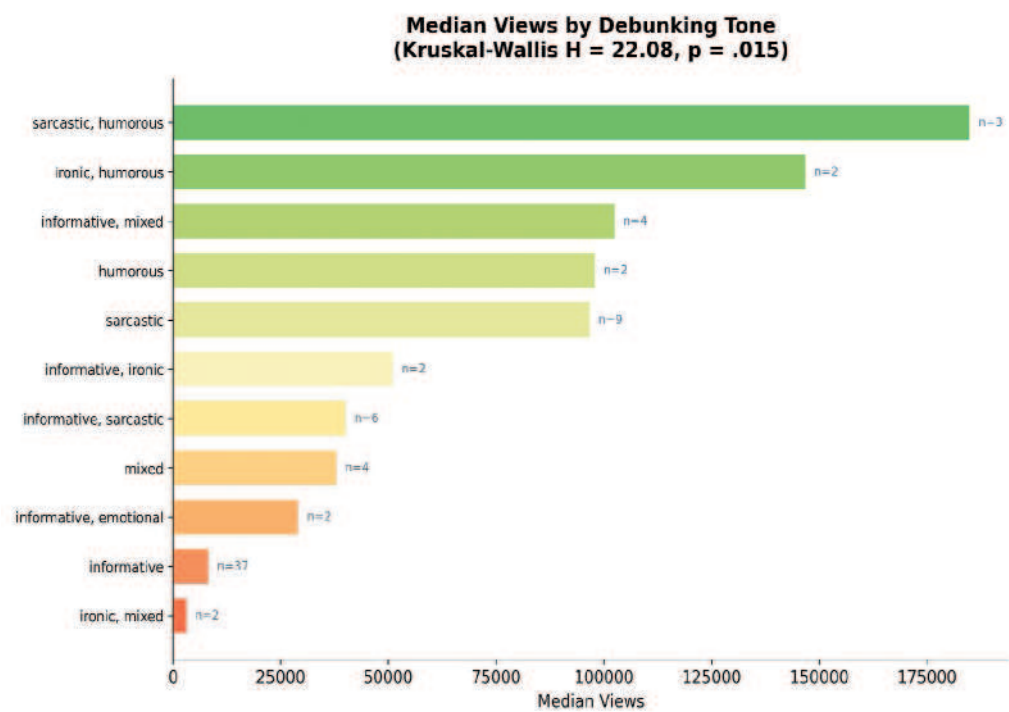
Analysis of structural characteristics revealed that the vast majority of debunking videos (96.2%, $n = 77$) were original stitches responding directly to misinformation content, rather than stitches of existing stitches. Only 3.8% ($n = 3$) of the corpus consisted of secondary stitches building upon prior corrective responses. This finding suggests that Romanian TikTok debunking operates primarily through distributed individual responses rather than sequential collaborative chains. From the perspective of participatory culture theory (Jenkins et al., 2016), this pattern may indicate that the algorithmic architecture of TikTok's For You feed, which surfaces content based on interest signals rather than social connections, reduces the incentive for users to build upon existing stitches when they can instead respond directly to the original misinformation source. For digital counter-misinformation practices, this distributed model implies that corrective saturation of a given claim may emerge organically through independent parallel responses rather than through coordinated chain formation.

4.3 RQ2: Which Rhetorical, Visual, and Tonal Strategies Are Associated with Higher Engagement?

4.3.1 Tone of Debunking Content and Engagement

Coding of the tonal register employed in debunking videos revealed a predominance of informative approaches (46.3%), though combinations incorporating sarcasm, irony, or humor were also prevalent. Statistical analysis revealed significant differences in viewership across tonal categories. A Kruskal-Wallis H test indicated that tone significantly predicted views ($H = 22.08, p = .015$). Figure 4 displays the relationship between tone and engagement.

Figure 4. Median views by debunking tone

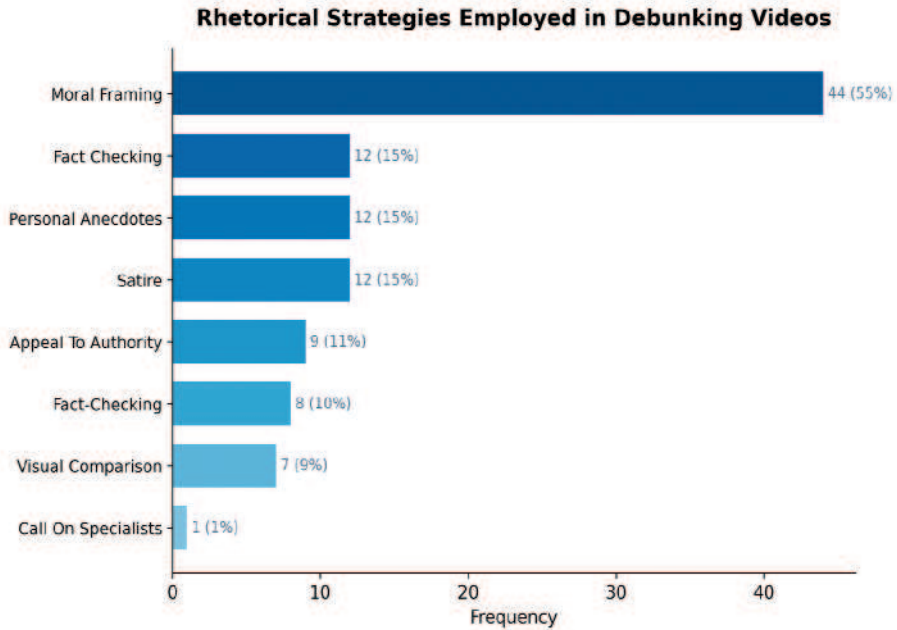


These findings indicate a tension between the prevalence of informative debunking strategies (46.3% of videos) and their relative engagement performance. While purely informative content dominated the corpus, it achieved median views of only 8,212 – substantially below the sarcastic + humorous combination (median = 184,900) and sarcastic content alone (median = 96,700).

4.3.2 Rhetorical Strategies

Content analysis of rhetorical approaches revealed moral framing as the dominant strategy employed by debunkers (Figure 5).

Figure 5. Rhetorical strategies employed in debunking videos

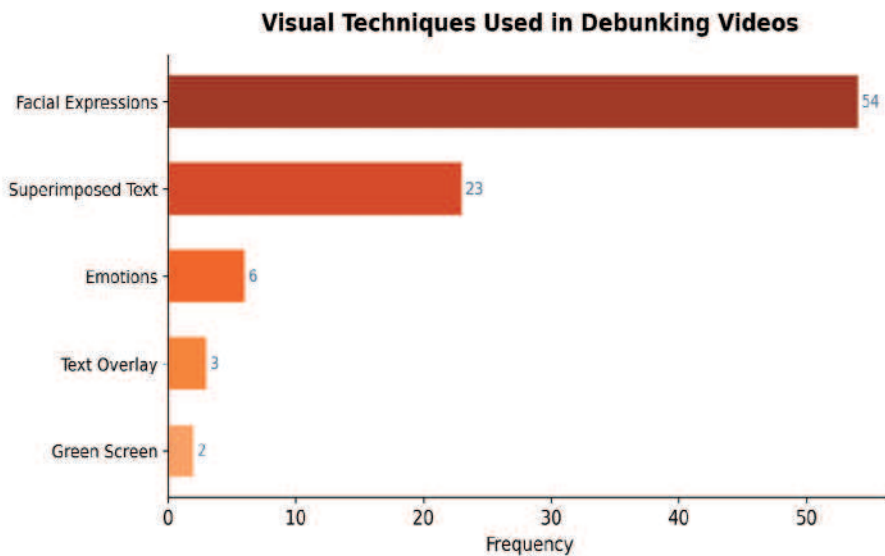


The prevalence of moral framing (41.9%) suggests that Romanian TikTok debunkers frequently position misinformation correction within ethical frameworks, appealing to values and normative standards rather than relying solely on factual verification. Fact-checking (19.0%) and personal anecdotes (11.4%) also featured prominently, indicating a hybrid approach combining evidence-based and experiential persuasion.

4.3.3 Visual Techniques

Analysis of visual elements revealed strong reliance on direct-to-camera address (Figure 6).

Figure 6. Visual techniques used in debunking videos

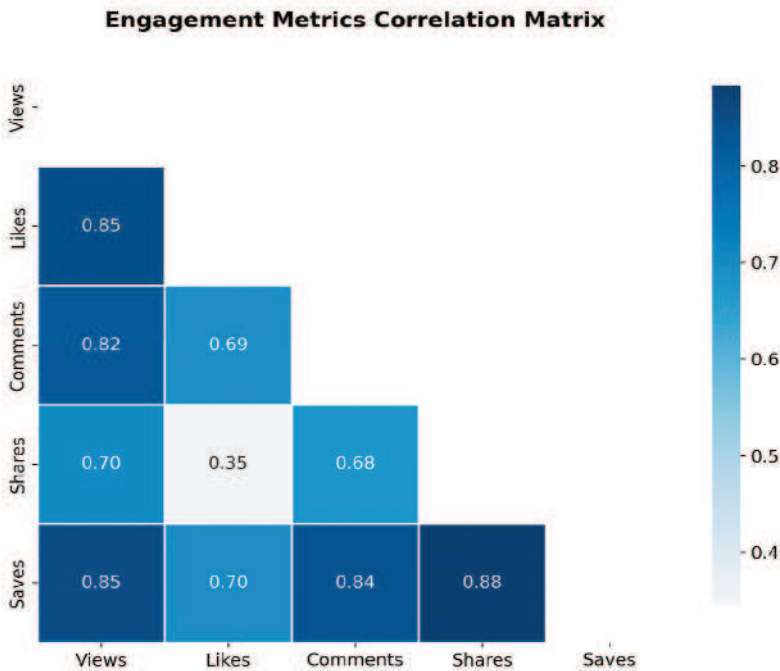


Facial expressions served as the primary visual element in 67.5% of videos, with superimposed text (28.8%) providing additional textual context. More technically demanding features such as green screen effects (2.5%) were comparatively rare, suggesting that stitch-based debunking relies primarily on personal address rather than elaborate production.

4.3.4 Engagement Correlations

Correlation analysis revealed strong relationships among engagement metrics (Figure 7). Follower count was significantly associated with views (Spearman’s $\rho = .457, p < .001$), while video duration showed no significant relationship with views ($\rho = .015, p = .89$).

Figure 7. Correlation matrix of engagement metrics

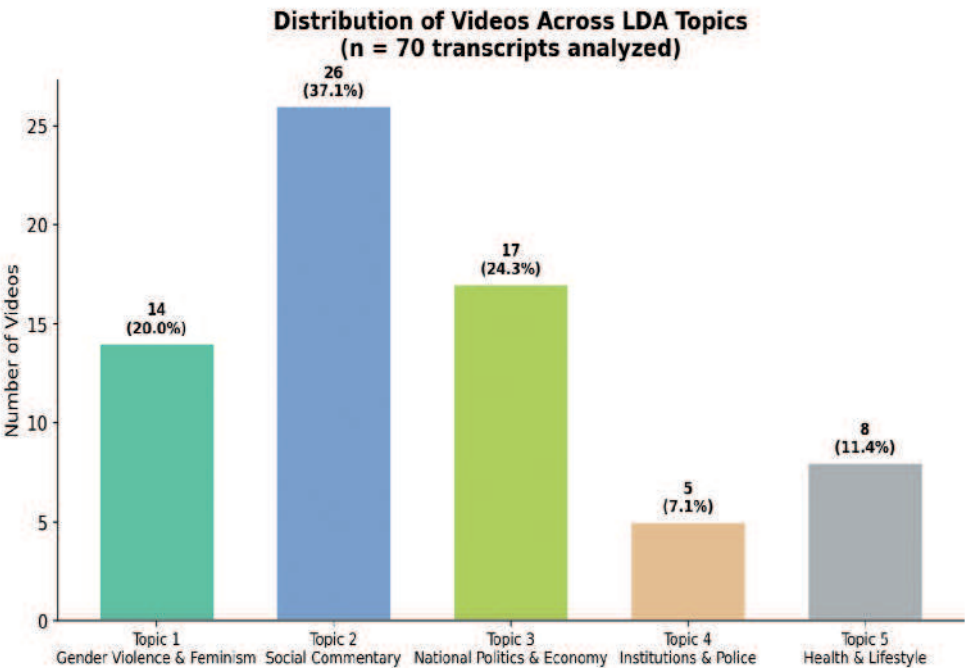


4.4 RQ3: How Does the Discourse Evolve Within Stitch Chains?

4.4.1 Topic Modeling of Transcript Content

To identify latent thematic structures in debunking discourse, Latent Dirichlet Allocation (LDA) topic modeling was applied to the transcribed content of 70 videos (10 videos had insufficient text for analysis). The model identified five distinct topical clusters (Figure 8).

Figure 8. Distribution of videos across LDA topics

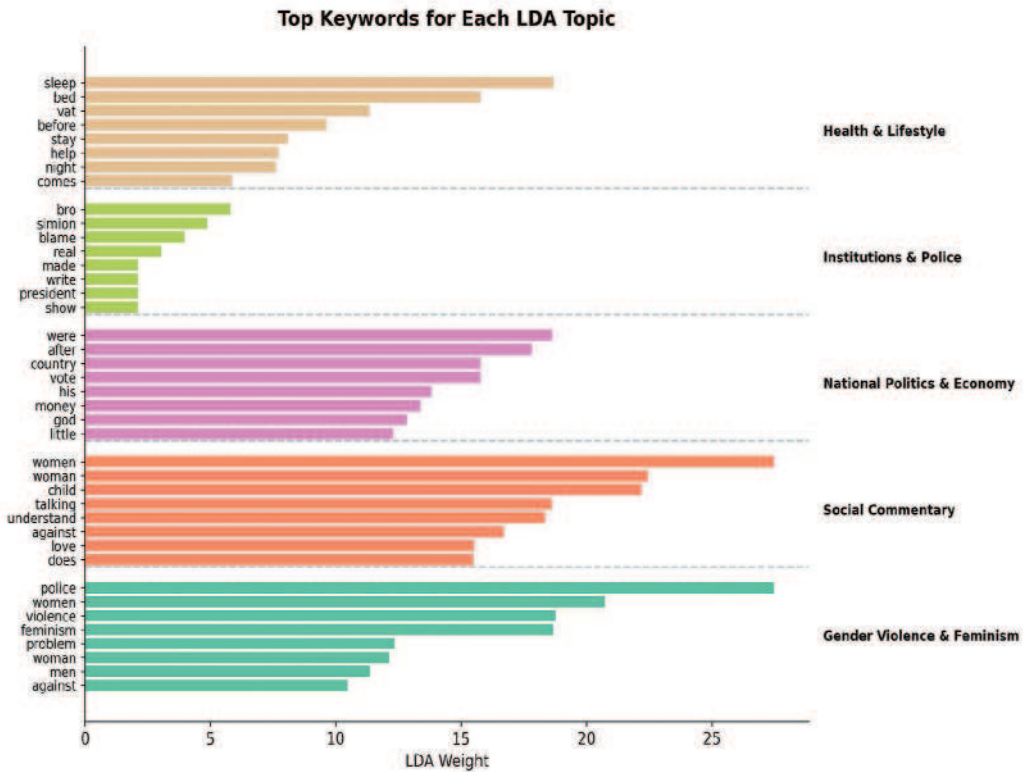


Topic 2 (Social Commentary) emerged as the largest cluster, encompassing 37.1% of transcripts and addressing discussions of women, children, and broader social issues. Topic 3 (National Politics & Economy) captured 24.3% of videos, focusing on Romania-specific concerns including electoral processes and economic policy. Topic 1 (Gender Violence & Feminism) explicitly addressed violence against women and feminist perspectives (20.0%).

4.4.2 Topic Keywords and Semantic Structure

Analysis of the top keywords for each topic reveals the semantic structure of debunking discourse (Figure 9).

Figure 9. Top keywords for each LDA topic

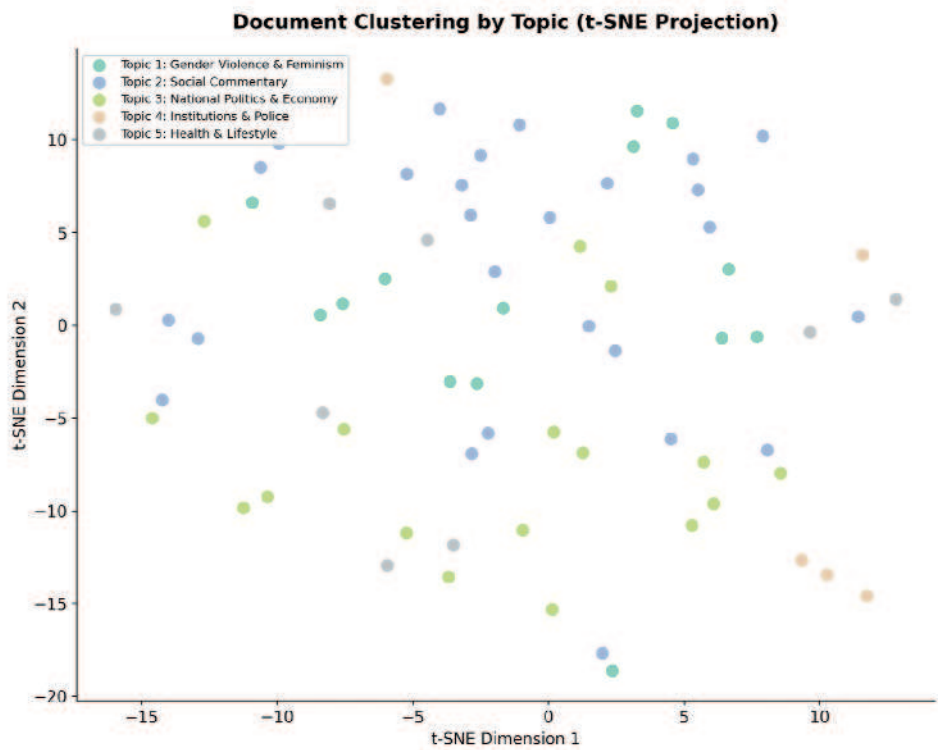


The keyword distributions reveal thematically coherent clusters. Topic 1 (Gender Violence & Feminism) is characterized by terms including ‘police,’ ‘women,’ ‘violence,’ and ‘feminism,’ reflecting discourse around gender-based violence and institutional responses. Topic 3 (National Politics & Economy) features ‘country,’ ‘vote,’ ‘money,’ and ‘god,’ indicating discussions of national identity, electoral processes, and economic concerns. Topic 5 (Health & Lifestyle) contains ‘sleep,’ ‘bed,’ and ‘vat,’ capturing health misinformation and lifestyle content.

4.4.3 Document Clustering

t-SNE dimensionality reduction of transcript content reveals the spatial relationships among documents (Figure 10).

Figure 10. Document clustering by topic (t-SNE projection)

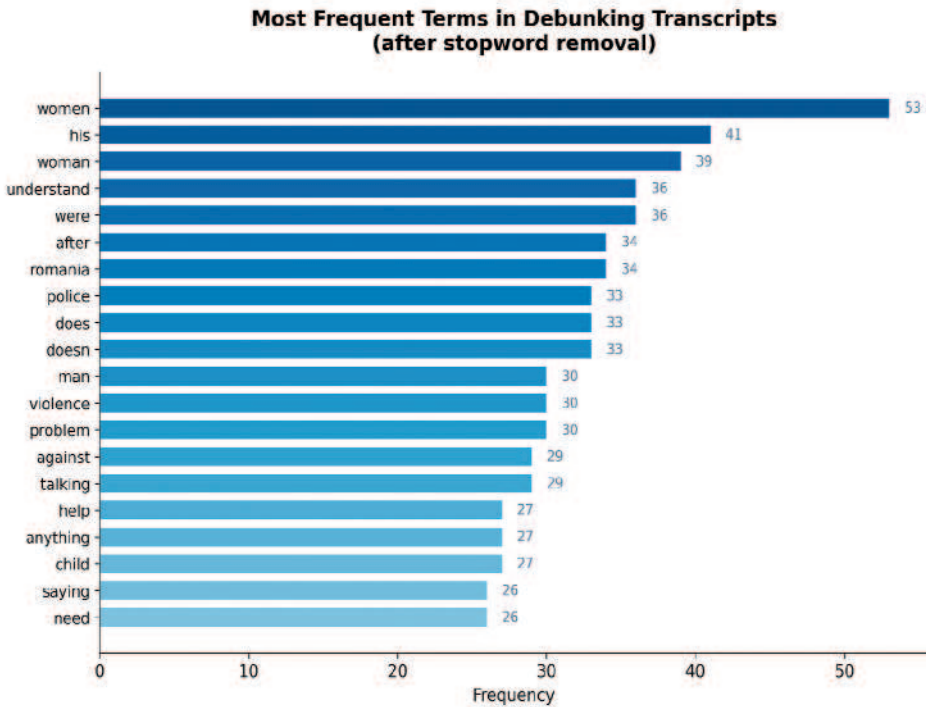


The visualization reveals moderate topical clustering with some overlap, suggesting that debunking videos often address multiple thematic concerns simultaneously. The distribution indicates that while distinct topical communities exist, debunkers frequently traverse multiple discourse domains.

4.4.4 Discourse Patterns

Linguistic analysis of transcripts revealed characteristic patterns in how debunkers construct corrective messages (Figure 11).

Figure 11. Most frequent terms in debunking transcripts



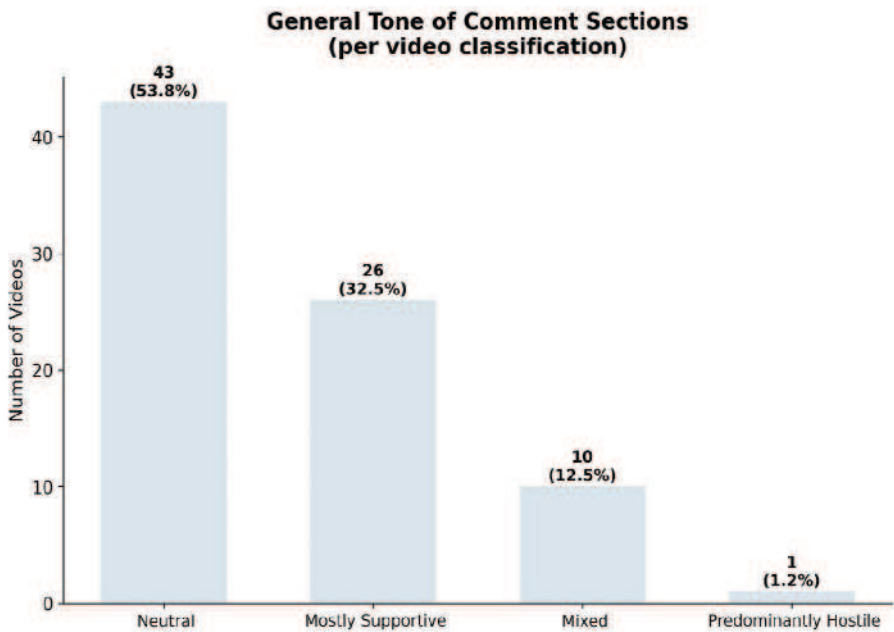
The predominance of terms related to women, gender, and social issues reflects the thematic focus identified in topic modeling. The frequent appearance of ‘understand’ and ‘fact’ suggests an emphasis on comprehension and evidence in debunking discourse.

4.5 RQ4: How Do TikTok Audiences Respond to Stitch-Based Debunks?

4.5.1 General Comment Tone

Analysis of the overall comment section tone for each video revealed predominantly neutral to supportive reception of debunking content (Figure 12).

Figure 12. General tone of comment sections (per video classification)

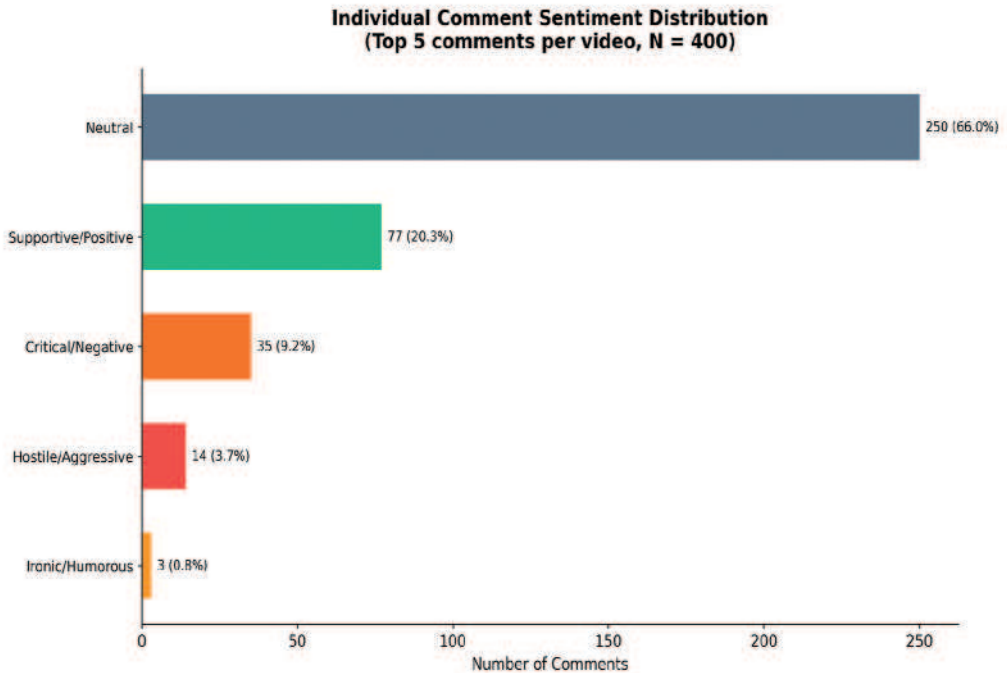


The majority of videos (53.8%) received neutral comment engagement, while 32.5% were characterized as mostly supportive. The rarity of predominantly hostile comment sections (1.3%) suggests that stitch-based debunking does not typically provoke substantial backlash from audiences.

4.5.2 Individual Comment Sentiment

Granular analysis of the top five comments per video (N = 400 comments) provided additional insight into audience response patterns (Figure 13).

Figure 13. Individual comment sentiment distribution

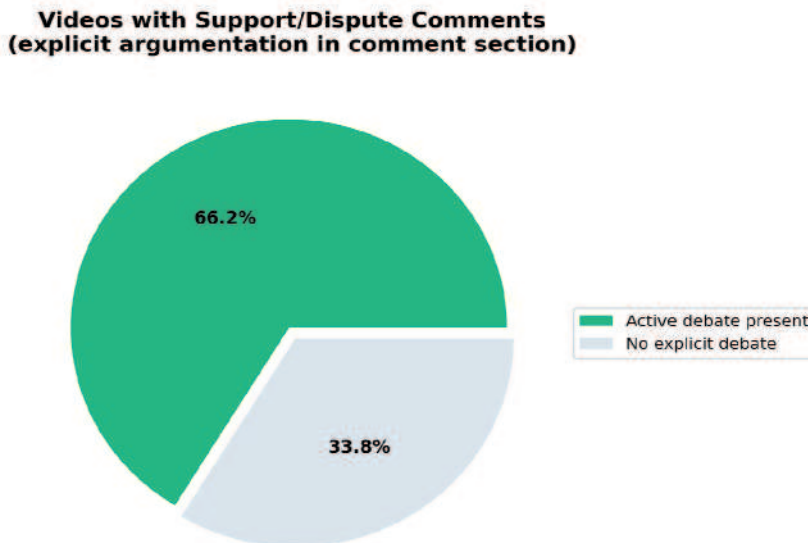


The distribution confirms the pattern observed at the video level: audiences predominantly respond with neutral engagement (66.0%) or explicit support (20.3%). Critical/negative comments represented only 9.2% of responses, while hostile/aggressive comments were rare (3.7%). This suggests that debunking content does not systematically provoke hostile counter-reactions.

4.5.3 Argumentation in Comments

Analysis of whether comments contained explicit support or dispute of the debunking claims revealed active engagement with the corrective content (Figure 14).

Figure 14 Presence of support/dispute comments in videos

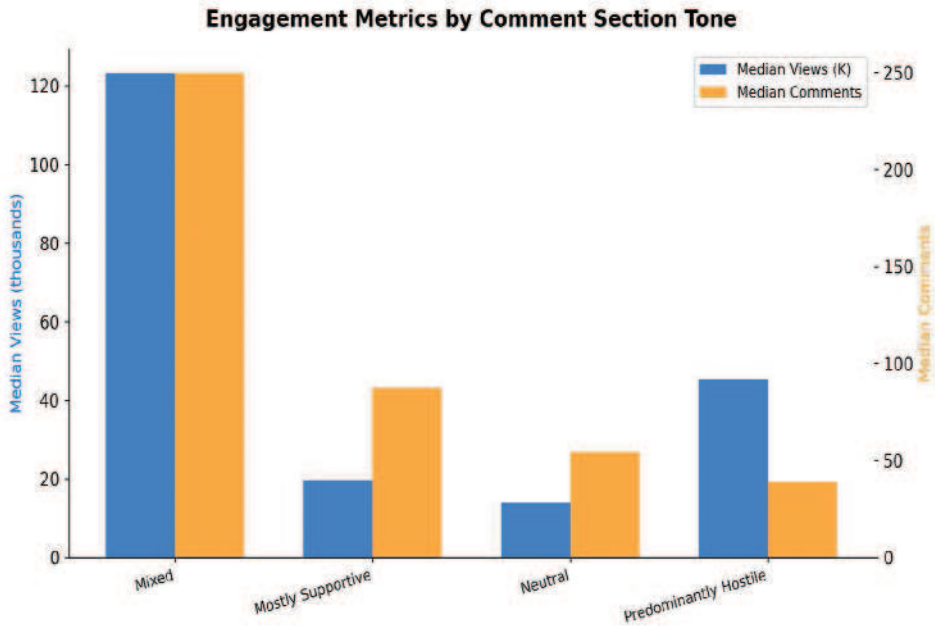


Most videos (66.2%, $n = 53$) contained comments that explicitly supported or disputed the debunk, indicating that audiences engage substantively with the factual claims presented rather than merely reacting to surface-level content.

4.5.4 Comment Tone and Engagement

Examination of the relationship between comment section tone and engagement metrics revealed interesting patterns (Figure 15).

Figure 15. Engagement metrics by comment section tone



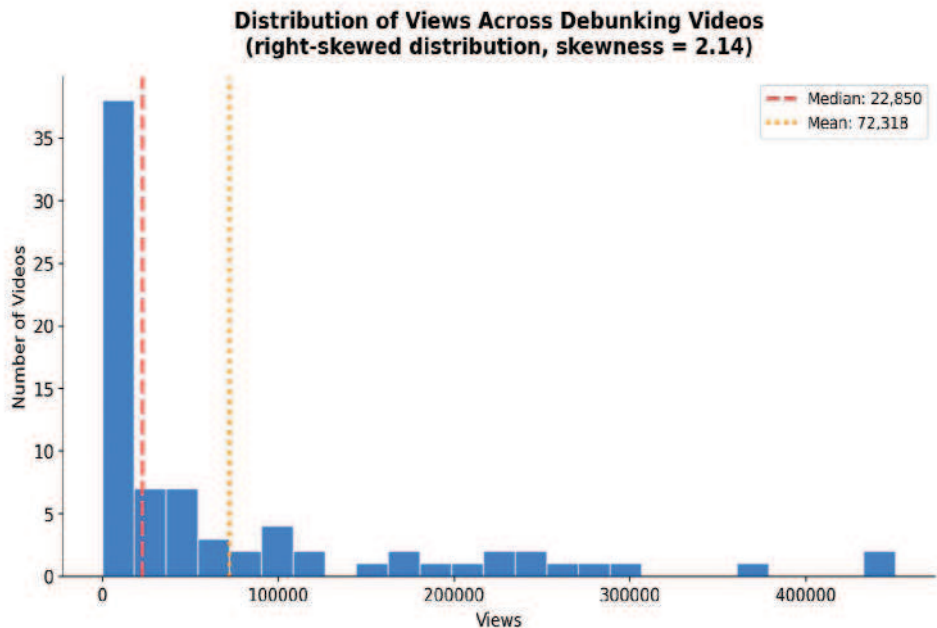
Videos with mixed comment sections showed higher median views than those with uniformly supportive or neutral responses, which could suggest that controversial debunking content may achieve greater algorithmic distribution through increased comment activity, though this pattern should be interpreted cautiously given the small sample size.

4.6 RQ5: To What Extent Do Stitch Chains Contribute to Viral Spread?

4.6.1 Views Distribution

The distribution of views exhibited strong positive skewness (skewness = 2.14), with a small number of highly successful videos substantially elevating the mean above the median (Figure 16).

Figure 16. Distribution of views across debunking videos

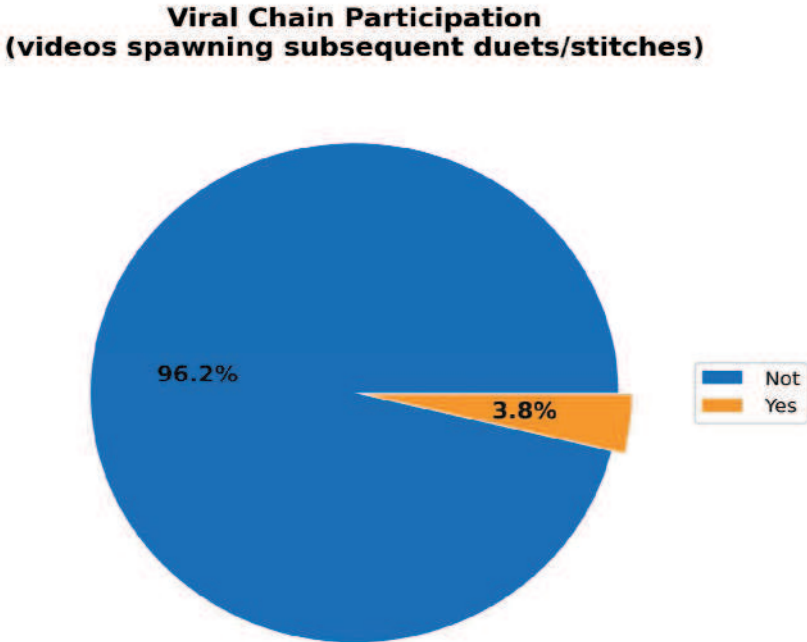


Despite the absence of chain amplification, individual debunking videos achieved substantial reach. The total corpus accumulated 5,640,788 views, with the top-performing video achieving 451,000 views as a standalone stitch.

4.6.2 Viral Chain Participation

Examination of whether debunking videos formed part of larger viral chains revealed limited chain formation (Figure 17).

Figure 17. Viral chain participation rates

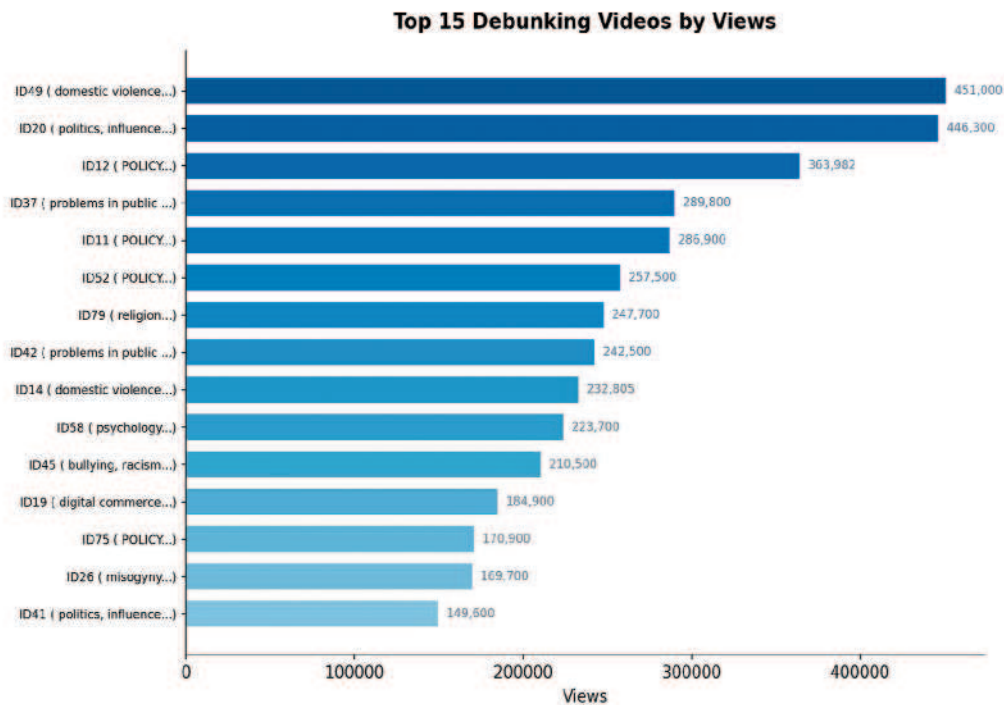


Only 3.8% ($n = 3$) of videos were identified as part of viral chains that spawned subsequent duets or stitches. The vast majority (96.2%) operated as standalone corrective interventions, suggesting that TikTok's algorithmic distribution, rather than user-driven chain formation, serves as the primary mechanism for amplifying debunking content.

4.6.3 Top Performing Content

Analysis of the highest-performing debunking videos reveals the topics and characteristics associated with viral reach (Figure 18).

Figure 18. Top 15 debunking videos by views



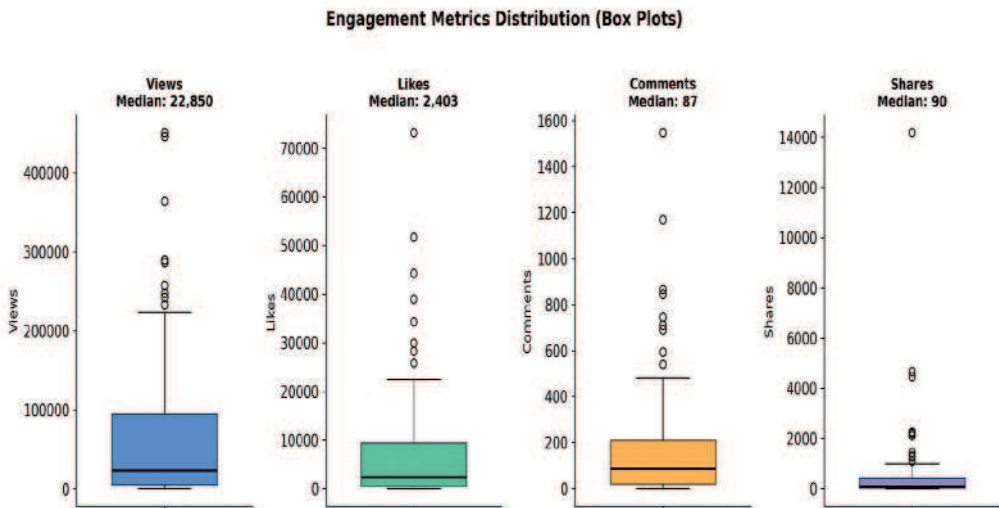
The top-performing videos addressed diverse topics including domestic violence, politics, public institutions, and psychology. This diversity suggests that viral potential is not confined to specific subject matter but depends on other factors including tone, timing, and creator following. We retain all 15 entries rather than a reduced subset because the full distribution reveals an analytically significant pattern: a steep drop-off separates the top three videos (363,982; 232,805; and 82,029 views respectively) from the remaining twelve, which cluster between approximately 5,800 and 14,600 views. Truncating to the top five would obscure this long-tail structure and overrepresent sarcastic content, which dominates the highest positions but is not representative of the broader corpus. Examining the five highest-performing videos in detail further illuminates the engagement-information tension identified in this study. The leading video (ID12, 363,982 views) addressed policy-related misinformation through sarcastic moral framing, while the second (ID14, 232,805 views) employed satire to debunk domestic violence narratives originating from a public artist. The third (ID13, 82,029 views) also used sarcastic satire, targeting protest-related misinformation from an influencer. Notably, the fourth-ranked video (ID10, 61,446 views) achieved substantial reach through an informative tone with moral framing on the same topic of domestic violence, suggesting that

informative content can achieve comparable visibility when paired with value-based appeals. The fifth (ID15, 14,616 views) combined a mixed tone with moral framing and visual comparison to address controversial topics. Collectively, the top five confirm that sarcastic and satirical approaches dominate viral reach, but also demonstrate that moral framing serves as a consistent rhetorical strategy across tonal registers, appearing in four of five top-performing videos regardless of whether the overall tone is sarcastic or informative.

4.6.4 Engagement Metrics Summary

Across all engagement metrics, the corpus demonstrated substantial audience interaction (Figure 19).

Figure 19. Engagement metrics distribution (box plots)



The engagement distributions reveal the characteristic right-skewed patterns of social media metrics, with median values substantially below means due to high-performing outliers. Over one-fifth of all videos (21.3%) exceeded 100,000 views, demonstrating that grassroots debunking can achieve substantial reach.

5 Discussion

This study examined stitch-based debunking on Romanian TikTok, offering an exploratory view of grassroots misinformation correction emerging from the intersection of platform affordances and participatory culture. The findings extend scholarship on networked debunking (Margolin et al., 2018) and civic media literacies (Mihailidis, 2018) by demonstrating

how platform-native features can be appropriated for epistemic purposes where institutional fact-checking remains underdeveloped. Rather than extending professional verification, stitch-based debunking constitutes a peer-to-peer corrective mechanism operating according to its own communicative logics.

The stitch feature's architectural properties prove consequential for understanding correction in short-form video environments. Unlike duets, which display content simultaneously, stitches create sequential juxtaposition mirroring the claim-then-correction structure recommended in debunking scholarship (Lewandowsky et al., 2020). This structure embeds corrective discourse directly within the same viewing experience as the misinformation, potentially circumventing separate publication channels characterizing institutional fact-checking (Graves, 2016). The five-second limitation forces creators to distill misleading claims to essential elements, potentially reducing risks of extended repetition that reinforces false information through continued influence effects (Lewandowsky et al., 2012). These findings align with Bucher and Helmond's (2018) framework distinguishing between low-level technical affordances and high-level communicative practices, suggesting TikTok's design creates material conditions for dialogic correction differing substantially from text-based platforms.

The predominance of peer-produced content as misinformation targets (57.5%) situates stitch-based debunking within participatory culture frameworks rather than traditional media criticism. This resonates with Bruns's (2018) concept of *gawatching*, yet stitch-based debunking extends it by adding an explicitly corrective dimension: creators actively contest and reframe content through audiovisual response (Bruns, 2018). The peer-to-peer character may confer advantages in perceived authenticity that institutional fact-checks lack (Vraga & Bode, 2017). Recent research supports this interpretation, finding peer corrections can be perceived as more trustworthy than expert-driven interventions (Chuai et al., 2024; Drolsbach et al., 2024). This aligns with Benesch and colleagues' (2016) distinction between organized counter-speech campaigns and organic community responses, with the latter potentially proving more persuasive precisely because they lack institutional sponsorship.

A noteworthy observation in this exploratory sample concerns the near-absence of extended stitch chains, with only 3.8% consisting of secondary stitches. Within the scope of this sample, this pattern is consistent with a model of distributed parallelism rather than sequential elaboration: multiple creators independently respond to common targets, producing cumulative corrective effects through volume rather than chain-based amplification. This diverges from networked content creation theorized by Abidin (2021), instead indicating that TikTok's algorithmic distribution, particularly the „For You“ feed's content-based recommendations (Zulli & Zulli, 2022), serves as the primary amplification vector. This aligns with research on human-algorithm interactions emphasizing how algorithmic amplification creates feedback loops affecting information spread (McLoughlin & Brady, 2024).

A notable pattern concerns the inverse relationship between prevalence and performance of tonal strategies. While purely informative debunking dominated (46.3%), it achieved substantially lower engagement than sarcastic or humorous content, with sarcastic-plus-humorous combinations achieving median views 22 times higher (184,900 vs. 8,212). This tension points to a challenge worth further investigation: strategies most likely to achieve algorithmic amplification may not be most effective at reducing misperceptions. Research has produced contested findings here. Boukes and Hameleers (2022) found satirical fact-checks were not uniformly more effective than regular corrections (Boukes & Hameleers, 2023). Vraga et al. (2019) demonstrated logic-based corrections were more credible among dismissive audi-

ences, while humor-based corrections proved more effective among those already convinced by false information (Vraga et al., 2019). Yeo et al. (2021) caution that humor may undermine source credibility depending on tone intensity (Yeo & McKasy, 2021). The prominence of moral framing (41.9%) illuminates how debunkers navigate this tension. By positioning correction within ethical frameworks, creators may achieve both engagement through value-based appeals and informational objectives through articulating normative standards. This resonates with Lakoff's (2018) emphasis on values in political communication, extending it to participatory fact-checking. The high frequency of direct address through facial expressions (67.5%) suggests debunkers leverage parasocial dynamics and perceived authenticity characteristic of TikTok's creator culture (Abidin, 2021). Experimental work supports this approach, finding citizen-led correction videos can effectively reduce belief in false claims when combining evidentiary content with engaging presentation (Bhargava et al., 2023).

Contrary to concerns about hostile backlash, audience response was predominantly neutral to supportive, with only 1.3% receiving hostile comment sections while 86.3% were neutral or supportive. This challenges assumptions about polarized responses and aligns with revisionist scholarship questioning backfire effect prevalence (Wood & Porter, 2019). The high rate of active engagement with factual claims (66.2% contained comments supporting or disputing the debunk) indicates audiences treat debunking as invitation to deliberation rather than provocation to resistance. This deliberative engagement represents distributed epistemic work, resonating with research showing that peer-based corrections foster collective claim evaluation (Drolsbach et al., 2024). These findings carry theoretical and practical implications. They suggest expanding fact-checking conceptions beyond institutional models to encompass grassroots practices operating according to different logics, balancing informational goals against engagement imperatives, echoing Bachmann and Valenzuela's (2023) observation that fact-checking effects depend on format and context (Bachmann & Valenzuela, 2023). The substantial cumulative reach (over 5.6 million views) demonstrates that platform-native correction represents a meaningful component of the information ecosystem. Media literacy programs might incorporate attention to platform-specific affordances, equipping citizens for collective verification processes. Research on algorithmic literacy supports this, finding that individuals who understand algorithmic functions are more likely to engage in corrective behaviors (Chung & Wihbey, 2024).

5.1 Limitations

Several limitations constrain generalizability. The purposive sampling strategy, while appropriate for exploratory research (Tavory & Timmermans, 2014), does not permit statistical generalization to broader TikTok debunking populations. The sample size (N=80) limited statistical power for subgroup analyses. Engagement metrics captured at a single time point may not reflect long-term patterns, and research suggests correction timing significantly affects impact (Chuai & Zhao, 2024). Analysis of only top five visible comments per video may obscure nuanced audience response patterns. The Romanian-language focus limits cross-cultural generalization, particularly given most existing TikTok misinformation research focuses on English-language content (Bhargava et al., 2023). Finally, the study examined practices rather than effectiveness; experimental research measuring attitude change would complement this descriptive approach, especially given meta-analytic findings showing correction effectiveness varies substantially across topics and audiences (Chan & Albarracín, 2023).

6 Conclusion

This exploratory study examined stitch-based debunking as a participatory fact-checking practice within the Romanian TikTok ecosystem. Findings reveal practices characterized by peer-to-peer correction, moral framing, multimodal persuasion, and predominantly supportive audience reception, patterns that distinguish platform-native debunking from institutional models while highlighting its complementary potential. The central tension between informative debunking prevalence and the superior engagement performance of humorous content illuminates challenges facing participatory correction in algorithmic environments. Creators must balance epistemic goals against platform imperatives rewarding emotional engagement and viral potential. Within this sample, moral framing appears to serve as one possible bridging strategy, though further research with larger and more representative samples would be needed to establish whether this pattern generalizes across TikTok debunking practices more broadly.

The near-absence of extended chains combined with substantial individual video reach suggests algorithmic distribution rather than user-driven chain formation serves as the primary amplification mechanism. For grassroots debunking, this means individual interventions can achieve substantial reach without collaborative chain formation, though visibility remains contingent on algorithmic dynamics creators cannot directly control. The predominantly supportive audience reception offers a counterpoint to polarization concerns, suggesting peer-to-peer debunking can unfold within deliberative frames. As misinformation challenges democratic discourse, understanding citizen self-organized corrective responses becomes urgent. Stitch-based debunking exemplifies how platform affordances can be appropriated for civic purposes, offering participatory media literacy models complementing institutional fact-checking. Future research should examine stitch-based debunking across contexts, track engagement longitudinally, experimentally assess engagement-optimized versus informative correction effectiveness, and investigate how algorithm changes affect corrective content circulation. The stitch feature, designed for creative remix, has been repurposed as an epistemic maintenance tool, illuminating both the potential of platform architecture and digital publics' participatory ingenuity navigating contemporary information disorder.

Author contributions

Eduard Gross: Conceptualization, Methodology, Formal Analysis, Investigation, Writing – Original Draft, Writing – Review & Editing. Anda-Ștefana DOBRE: Data Curation, Investigation, Writing – Review & Editing.

AI authorship and AI use

No generative AI tools were used to produce substantive content for this manuscript. Any use of AI was limited to routine copy-editing (improving readability, grammar, and formatting), with final responsibility for all content resting with the authors.

Ethics declarations

Ethical approval was not required as no human participants were involved in this research. The study analysed publicly available content posted on TikTok.

Competing interests

The authors declare no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data availability

The data supporting the findings of this study were collected from publicly available content on TikTok. Owing to platform terms of service and the potential to identify content creators, the dataset is not publicly archived but is available from the corresponding author on reasonable request.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Abidin, C. (2020). Mapping Internet Celebrity on TikTok: Exploring Attention Economies and Visibility Labours. *Cultural Science Journal*, 12(1), 77–103. <https://doi.org/10.5334/csci.140>
- Aral, S. (2021). *The hype machine: How social media disrupts our elections, our economy, and our health – and how we must adapt* (First edition). Currency.
- Bachmann, I., & Valenzuela, S. (2023). Studying the Downstream Effects of Fact-Checking on Social Media: Experiments on Correction Formats, Belief Accuracy, and Media Trust. *Social Media + Society*, 9(2), 20563051231179694. <https://doi.org/10.1177/20563051231179694>
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132. <https://doi.org/10.1126/science.aaa1160>
- Benesch, S. (2016). *Counterspeech on Twitter: A Field Study*. <https://www.dangerousspeech.org/libraries/counterspeech-on-twitter-a-field-study>
- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (1st ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780190923624.001.0001>
- Berger, J. (2016). *Contagious: Why things catch on* (First Simon & Schuster trade paperback edition). Simon & Schuster Paperbacks.
- Bhargava, P., MacDonald, K., Newton, C., Lin, H., & Pennycook, G. (2023). How effective are TikTok misinformation debunking videos? *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-114>

- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(null), 993–1022.
- Boukes, M., & Hameleers, M. (2023). Fighting lies with facts or humor: Comparing the effectiveness of satirical and regular fact-checks in response to misinformation and disinformation. *Communication Monographs*, 90(1), 69–91. <https://doi.org/10.1080/03637751.2022.2097284>
- Brewer, P. R., & McKnight, J. (2017). “A Statistically Representative Climate Change Debate”: Satirical Television News, Scientific Consensus, and Public Perceptions of Global Warming. *Atlantic Journal of Communication*, 25(3), 166–180. <https://doi.org/10.1080/15456870.2017.1324453>
- Bruns, A. (2009a). *Blogs, Wikipedia, Second Life, and beyond: From production to produsage*. Lang.
- Bruns, A. (2009b). *Gatewatching: Collaborative online news production* (Nachdr.). Lang.
- Bruns, A. (2018). *Gatewatching and News Curation*. Peter Lang US. <https://doi.org/10.3726/b13293>
- Bucher, T., & Helmond, A. (2018). The Affordances of Social Media Platforms. In J. Burgess, A. Marwick, & T. Poell, *The SAGE Handbook of Social Media* (pp. 233–253). SAGE Publications Ltd. <https://doi.org/10.4135/9781473984066.n14>
- Chan, M. S., & Albarracín, D. (2023). A meta-analysis of correction effects in science-relevant misinformation. *Nature Human Behaviour*, 7(9), 1514–1525. <https://doi.org/10.1038/s41562-023-01623-8>
- Chuai, Y., Pilarski, M., Renault, T., Restrepo-Amariles, D., Troussel-Clément, A., Lenzini, G., & Pröllochs, N. (2024). *Community-based fact-checking reduces the spread of misleading posts on social media* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2409.08781>
- Chung, M., & Wibbey, J. (2024). The algorithmic knowledge gap within and between countries: Implications for combatting misinformation. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-155>
- Corbu, N., Oprea, D.-A., Negrea-Busioc, E., & Radu, L. (2020). ‘They can’t fool me, but they can fool the others!’ Third person effect and fake news detection. *European Journal of Communication*, 35(2), 165–180. <https://doi.org/10.1177/0267323120903686>
- Cușnir, C. (2025). TikTok Design and Experience: A Semiotic Technology Approach. *Romanian Journal of Communication and Public Relations*, 27(1), 7–23. <https://doi.org/10.21018/rjcp.2025.1.683>
- Davis, J. L., & Chouinard, J. B. (2016). Theorizing Affordances: From Request to Refuse. *Bulletin of Science, Technology & Society*, 36(4), 241–248. <https://doi.org/10.1177/0270467617714944>
- Dumitrescu, R. (2023, October 16). *European Commission in Romania launches anti-disinformation platform*. Romania Insider. <https://www.romania-insider.com/european-commission-romania-anti-disinformation-platform-2023>
- Drolsbach, C. P., Solovev, K., & Pröllochs, N. (2024). Community notes increase trust in fact-checking on social media. *PNAS Nexus*, 3(7), pgae217. <https://doi.org/10.1093/pnasnexus/pgae217>
- Gamson, W. A., & Modigliani, A. (1989). Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach. *American Journal of Sociology*, 95(1), 1–37. <https://doi.org/10.1086/229213>
- Gibson, J. J. (2015). *The ecological approach to visual perception: Classic edition*. Psychology Press.
- Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press. <https://doi.org/10.12987/9780300235029>
- Graves, L. (2016). *Deciding what’s true: The rise of political fact-checking in American journalism*. Columbia University Press.
- Gray, J., Jones, J. P., & Thompson, E. (2009). *Satire TV: Politics and comedy in the post-network era*. NYU Press.
- Hameleers, M., Bos, L., Fawzi, N., Reinemann, C., Andreadis, I., Corbu, N., Scherer, C., Schulz, A., Shaefer, T., Aalberg, T., Axelsson, S., Berganza, R., Cremonesi, C., Dahlberg, S., De Vreese, C. H., Hess, A., Kartounidou, E., Kasproicz, D., Matthes, J., ... Weiss-Yaniv, N. (2018). Start Spreading the News: A Comparative Experiment on the Effects of Populist Communication on Political Engagement in Sixteen European Countries. *The International Journal of Press/Politics*, 23(4), 517–538. <https://doi.org/10.1177/1940161218786786>
- Iyengar, S., & Hahn, K. S. (2009). Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use. *Journal of Communication*, 59(1), 19–39. <https://doi.org/10.1111/j.1460-2466.2008.01402.x>
- Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. New York University Press.

- Jenkins, H., Ito, M., & Boyd, D. (2016). *Participatory culture in a networked era: A conversation on youth, learning, commerce, and politics*. Polity Press.
- Lakoff, G., Dean, H., & Hazen, D. (2004). *Don't think of an elephant! Know your values and frame the debate?; the essential guide for progressives*. Chelsea Green Publishing.
- Lelkes, Y. (2016). Mass Polarization: Manifestations and Measurements. *Public Opinion Quarterly*, 80(S1), 392–410. <https://doi.org/10.1093/poq/nfw005>
- Lewandowsky, S., Cook, J., & Lombardi, D. (2020). *Debunking Handbook 2020* [Dataset]. Databrary. <https://doi.org/10.17910/B7.1182>
- Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the “post-truth” era. *Journal of Applied Research in Memory and Cognition*, 6(4), 353–369. <https://doi.org/10.1016/j.jarmac.2017.07.008>
- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131. <https://doi.org/10.1177/1529100612451018>
- Margolin, D. B., Hannak, A., & Weber, I. (2018). Political Fact-Checking on Twitter: When Do Corrections Have an Effect? *Political Communication*, 35(2), 196–219. <https://doi.org/10.1080/10584609.2017.1334018>
- Marwick, A., Pippert, C., Furl, K., & Schnabel, E. (2025). Shapeshifters and Starseeds: Populist Knowledge Production, Generous Epistemology, and Disinformation on U.S. Conspiracy TikTok. *Social Media + Society*, 11(3), 20563051251357483. <https://doi.org/10.1177/20563051251357483>
- McLoughlin, K. L., & Brady, W. J. (2024). Human-algorithm interactions help explain the spread of misinformation. *Current Opinion in Psychology*, 56, 101770. <https://doi.org/10.1016/j.copsyc.2023.101770>
- Mihailidis, P. (2018). Civic media literacies: Re-Imagining engagement for civic intentionality. *Learning, Media and Technology*, 43(2), 152–164. <https://doi.org/10.1080/17439884.2018.1428623>
- Newman, N., Arguedas, my R., Robertson, C., Nielsen, R. K., & Fletcher, R. (2025, June 17). *Digital News Report 2025* | Reuters Institute for the Study of Journalism. <http://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025>
- Nyhan, B., & Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior*, 32(2), 303–330. <https://doi.org/10.1007/s11109-010-9112-2>
- Papacharissi, Z. (2015). *Affective publics: Sentiment, technology, and politics*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199999736.001.0001>
- Pariser, E. (2012). *The filter bubble: What the Internet is hiding from you*. Penguin Books.
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50. <https://doi.org/10.1016/j.cognition.2018.06.011>
- Pilipets, E., & Chao, J. (2025). Noise in sonic social media: Memetic soundscapes of Deep TikTok. *New Media & Society*, 14614448251358752. <https://doi.org/10.1177/14614448251358752>
- Shifman, L. (2014). *Memes in digital culture*. MIT Press.
- Stancea, A., Călin, I., & Ciocîrlan, C. (2025). Memes and Their Role in the 2019 Romanian Presidential Elections: An Exploratory Analysis. *Romanian Journal of Communication and Public Relations*, 26(2), 7–24. <https://doi.org/10.21018/rjcp.2024.2.593>
- Sunstein, C. R. (2017). *#Republic: Divided democracy in the age of social media*. Princeton university press.
- Tandoc, E. C., Lim, D., & Ling, R. (2020). Diffusion of disinformation: How social media users respond to fake news and why. *Journalism*, 21(3), 381–398. <https://doi.org/10.1177/1464884919868325>
- Tavory, I., & Timmermans, S. (2014). *Abductive analysis: Theorizing qualitative research*. The University of Chicago press.
- Thorson, E. (2016). Belief Echoes: The Persistent Effects of Corrected Misinformation. *Political Communication*, 33(3), 460–480. <https://doi.org/10.1080/10584609.2015.1102187>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Vraga, E. K., & Bode, L. (2017). Using Expert Sources to Correct Health Misinformation in Social Media. *Science Communication*, 39(5), 621–645. <https://doi.org/10.1177/1075547017731776>

Vraga, E. K., Kim, S. C., & Cook, J. (2019). Testing Logic-based and Humor-based Corrections for Science, Health, and Political Misinformation on Social Media. *Journal of Broadcasting & Electronic Media*, 63(3), 393–414. <https://doi.org/10.1080/08838151.2019.1653102>

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Council of Europe. <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>

Wood, T., & Porter, E. (2019). The Elusive Backfire Effect: Mass Attitudes’ Steadfast Factual Adherence. *Political Behavior*, 41(1), 135–163. <https://doi.org/10.1007/s11109-018-9443-y>

Yeo, S. K., & McKasy, M. (2021). Emotion and humor as misinformation antidotes. *Proceedings of the National Academy of Sciences*, 118(15), e2002484118. <https://doi.org/10.1073/pnas.2002484118>

Zulli, D., & Zulli, D. J. (2022). Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media & Society*, 24(8), 1872–1890. <https://doi.org/10.1177/1461444820983603>

Appendix A: Codebook for Systematic Content Analysis

This codebook operationalizes the 37 variables used for systematic content analysis of the video corpus. Variables are organized across eight analytical dimensions. For nominal variables permitting multiple codes, combinations are separated by commas in the dataset. Categories for rhetorical strategies were derived deductively from established typologies in misinformation correction research (Lewandowsky et al., 2020; Vraga et al., 2019) and refined inductively through iterative pilot coding of a subsample. Inter-coder reliability was assessed on a subset of 15 videos, achieving acceptable agreement levels across primary coding categories.

A. Video Metadata

Variable	Type	Values / Categories	Operational Definition
VIDEO_ID	Nominal	Unique identifier (ID01–ID80)	Sequential identifier assigned to each video in the corpus.
URL	String	TikTok URL	Direct link to the TikTok video.
Upload date	Date	DD/MM/YYYY	Date the video was published on TikTok.
Video duration	Continuous	Seconds	Total length of the video in seconds.
Username	Nominal	TikTok handle	Creator’s TikTok username at time of collection.
Number of followers	Continuous	Integer	Creator’s follower count at time of data collection.

B. Content & Structure

Variable	Type	Values / Categories	Operational Definition
Is it a remix/stitch of another stitch?	Binary	Yes / No	Whether the video is a secondary stitch (i.e., a stitch of an existing stitch) or an original stitch responding directly to the misinformation source.
Topic of disinformation targeted	Nominal (multi-code)	politics; health; domestic violence; feminism; misogyny; psychology; economy; bullying; religion; current affairs; legislation; justice; digital commerce; artificial intelligence; protests; animal neglect; popular myths; problems in public institutions; human behaviors; dependence; abuse; controversial topics; influencer; policy	Thematic category of the misinformation claim being debunked. Multiple codes permitted when a video addresses overlapping topics (e.g., 'bullying, education'). Categories were derived inductively from iterative viewing of the corpus.
Origin of the initial misinformation	Nominal	Content creators; Influencer; Government official (parliamentarian, politician, member of government); Journalist; TV station; Artist	Source type of the original misinformation being stitched. 'Content creators' refers to users who produce original content without a verified public profile or institutional affiliation. 'Influencer' refers to users who maintain established followings and monetized presences (e.g., brand deals, promotional content).

C. Tone & Rhetorical Strategies

Variable	Type	Values / Categories	Operational Definition
Tone of the debunk	Nominal (multi-code)	Informative; Sarcastic; Ironic; Humorous; Emotional; Mixed	Tonal register of the debunking response. Informative: neutral, evidence-based delivery. Sarcastic: sharp, contemptuous commentary intended to mock. Ironic: statements whose intended meaning differs from literal meaning. Humorous: lighthearted, comedic framing. Emotional: appeals to feeling, outrage, or empathy. Mixed: combinations not captured by other codes. Multiple codes permitted (e.g., 'informative, sarcastic').
Rhetorical strategies	Nominal (multi-code)	Moral framing; Fact-checking; Personal anecdotes; Appeal to authority; Satire; Visual comparison; Call on specialists	Persuasive strategies employed in the debunking message. Moral framing: positions correction within ethical or normative arguments appealing to shared values. Fact-checking: presents verifiable evidence, data, or sources to contradict false claims. Personal anecdotes: draws on personal experience or testimony as counter-evidence. Appeal to authority: invokes professional expertise, credentials, or institutional credibility. Satire: uses sustained comedic imitation to critique the original claim. Visual comparison: employs side-by-side or sequential visual evidence. Call on specialists: explicitly invites or references domain experts. Categories derived deductively from Lewandowsky et al. (2020) and Vraga et al. (2019), refined inductively through pilot coding.
Visual techniques	Nominal (multi-code)	Facial expressions; Superimposed text / text overlay; Green screen; Emotions	Primary visual elements employed. Facial expressions: direct-to-camera address where the creator's face and expressions are the main visual element. Superimposed text: on-screen text overlays providing context, emphasis, or captions. Green screen: use of TikTok's green screen effect to display background images/videos. Emotions: exaggerated or performative emotional display as a visual strategy.
Sound elements	Nominal	Original sound; Added music; TikTok trending sound	Audio track used. Original sound: creator's own voice/audio. Added music: non-trending background music selected by creator. TikTok trending sound: use of a currently popular sound on the platform.

D. Creator Positioning

Variable	Type	Values / Categories	Operational Definition
Number of similar clips per profile	Continuous	Integer	Count of debunking-related stitch videos on the same creator's profile, indicating systematic vs. occasional debunking engagement.
Bio suggests social/political involvement?	Binary	Yes / No	Whether the creator's TikTok bio text indicates affiliation with social causes, political organizations, NGOs, or civic initiatives.
Signs of influencer behavior	Nominal	Yes / No / Hazy	Whether the creator's profile shows evidence of monetization (brand deals, promotions, affiliate links). 'Hazy' indicates ambiguous cases.
Interpretive motivation	Nominal (multi-code)	Civic engagement; Identity performance; Audience search; Humor	Inferred primary motivation for producing the debunking video. Civic engagement: creator appears motivated by public interest or corrective intent. Identity performance: debunking serves self-presentation or persona construction. Audience search: content appears oriented toward maximizing reach/engagement. Humor: primary motivation appears to be entertainment rather than correction.

E. Engagement Metrics

Variable	Type	Values / Categories	Operational Definition
Number of Likes	Continuous	Integer	Like count at time of data collection.
Number of Comments	Continuous	Integer	Comment count at time of data collection.
Number of Distributions	Continuous	Integer	Share/distribution count at time of data collection.
Number of Views	Continuous	Integer	View count at time of data collection.
Number of Saves	Continuous	Integer	Save/bookmark count at time of data collection.

F. Virality & Chain Structure

Variable	Type	Values / Categories	Operational Definition
Part of a viral chain?	Binary	Yes / No	Whether the video spawned subsequent duets or stitches forming a chain of responses.

G. Audience Response

Variable	Type	Values / Categories	Operational Definition
Comments 1–5	String	Verbatim text	The top five publicly visible comments at time of collection, recorded verbatim.
Comment sentiment (per comment)	Nominal	Positive; Supportive; Neutral; Negative; Hostile; Ironic; Ironic/humorous	Affective valence of each individual comment. Positive: expresses agreement or approval. Supportive: explicitly endorses the debunking claim. Neutral: factual, off-topic, or emotionally unmarked. Negative: expresses disagreement or criticism. Hostile: aggressive, insulting, or threatening language. Ironic: sarcastic commentary whose intended meaning differs from literal.
Support/dispute of debunk in comments?	Binary	Yes / No	Whether any of the top five comments explicitly support or dispute the factual claims made in the debunking video.
General tone of comments	Nominal	Mostly supportive; Neutral; Mixed; Predominantly hostile	Overall characterization of the comment section's affective orientation. Coded holistically based on the balance of sentiment across the top five comments.

H. Captions & Hashtags

Variable	Type	Values / Categories	Operational Definition
Captions / Hashtags	String	Free text	Video caption text and hashtags used by the creator, recorded verbatim.