

TikTok Comedy During COVID-19: A Multimodal Study of Platform Data, Audience Engagement, and Creative Performance

Dan Bai

A thesis submitted for the degree of Doctor of Philosophy

(Applied Social and Economic Research)

Institute For Social and Economic Research (ISER)

University of Essex

October 2026

Abstract

This thesis investigates how multimodal features, platform infrastructure, and creative strategy jointly shape audience engagement in TikTok comedy content produced during the COVID-19 period (2020 to 2022). Drawing on a shared corpus of 741 videos from 14 high-profile creators active in the United Kingdom and the United States, three interrelated empirical studies address this aim from complementary analytical angles: evaluating the data infrastructure through which platform content is accessed, modelling how visual, audio, linguistic, and platform features jointly predict audience engagement, and examining how a specific creative strategy independently predicts engagement above and beyond correlated production characteristics. The three studies form a progressive analytical sequence in which each paper establishes the methodological and conceptual foundation for the next.

The first paper subjects TikTok's Research API to the first systematic comparative evaluation of its kind, contrasting its outputs against a commercial third-party analytics service and manual data collection across a broader corpus of 6,373 videos. The API offers genuine strengths, including scalability and access to 22 variables, eight of which are unavailable through alternative methods. The limitations are substantial: transcription coverage reaches approximately ten per cent of content, with systematic demographic skews favouring American male creators, and engagement metrics are inconsistent across data sources. The paper advocates multi-method auditing protocols as a necessary condition of validity for platform-based social science research.

The second paper examines how visual composition, audio context, linguistic markers, and platform affordances jointly predict three engagement outcomes in TikTok comedy videos, using Generalised Additive Mixed Models with creator-specific random intercepts. Likes, comments, and shares respond to qualitatively distinct multimodal configurations rather than a single underlying engagement factor. Likes function as sensory verdicts, sensitive to visual composition and prosodic expressiveness. Comments reflect cognitive verdicts, driven by video duration, title sentiment, and linguistic complexity. Shares embody social verdicts, shaped by relationship-focused content, transcript sentiment, and lexical accessibility. Video description length is the sole universal predictor across all three models. Creator gender consistently moderates multimodal effects throughout.

The third paper investigates role-play, a comedic strategy in which creators explicitly embody characters distinct from their everyday personas, accounting for 57.4 per cent of the corpus. Role-play videos display a consistent and identifiable production signature, characterised by significantly higher rates of accessory use, distinct costuming, and relationship-focused narratives. Role-play independently predicts approximately 21 per cent higher like counts and 42 to 48 per cent higher share counts after controlling for the full production feature set, yet shows no significant independent effect on comment counts. Male creators derive substantially greater sharing gains from same-gender role-play than female creators, a pattern consistent with accounts of gender-incongruent performance in algorithmically mediated contexts.

Across the three papers, two findings recur with sufficient consistency to constitute thesis-level contributions. Engagement on TikTok is irreducibly multidimensional: likes, comments, and shares are driven by distinct feature configurations and cannot be collapsed into a single outcome measure without substantial loss of analytical precision. Creator gender is a structural moderating variable that shapes not only baseline engagement levels but the mechanisms through which multimodal features and creative strategies translate into audience response. Together, the three studies advance multimodal communication theory, extend performance and gender performativity frameworks into short-form digital contexts, and contribute to an emerging methodological literature on the conditions of validity for platform-based social science.

Keywords: TikTok, short-form video, comedy, COVID-19, multimodal engagement, platform data, role-play, gender, algorithmic mediation, computational social science

Code and Data Availability Statement

The analytical code and processed dataset associated with this thesis will be made publicly available on the Open Science Framework (OSF) and/or GitHub following final examination and any post-viva revisions. Links and repository identifiers will be added to this page upon deposit. Readers wishing to request access to the code prior to public release may contact the author directly.

The raw video data cannot be shared publicly owing to TikTok's Terms of Service and applicable data protection regulations. Engagement metrics and derived feature data will be shared to the extent permitted by those constraints.

Acknowledgements

This thesis would not exist without the generosity of two people, and I want to say that plainly before anything else.

To Cara: thank you for taking me on. You came into this project at a moment when I was not in a good place, and you did so without hesitation and without condition. The steadiness and patience you brought to our work together made it possible for me to finish what I had started. I am more grateful than I know how to say.

To Yan: where do I begin. We first met in a psychology lab, monitoring an exam together, which is perhaps not the most obvious starting point for a research partnership. But from that beginning you have given me something I did not always feel I deserved: your trust. Your feedback has been constructive in the truest sense of the word, building rather than simply correcting, and your confidence in this project has carried me through the stretches where my own faltered. I have been very lucky to have you, and I know it.

Abstract i

Code and Data Availability Statement iii

Acknowledgements v

Introduction 7

Chapter 2: Harnessing Big Data, Hindered by Bias: Evaluating TikTok Research API for Fair and Optimal Social Sciences 38

Chapter 3: Sensory, Cognitive, and Social Verdicts: Multimodal Predictors of Likes, Comments, and Shares in TikTok Comedy 92

Chapter 4: Character Performance and Engagement: Multimodal Production Elements and Gender Effects in TikTok Role-Play Comedy 145

Conclusion 212

Introduction

1. TikTok as a Platform for Studying Social Phenomena: Digital Data, Comedic Content, and the COVID-19 Context

Social media platforms have become primary archives for the study of contemporary social phenomena. TikTok, in particular, generated a distinctive and large-scale record of cultural life during the COVID-19 pandemic. Between 2020 and 2022, as governments imposed lockdowns that disrupted livelihoods and everyday routines (Lyon & Coleman, 2023), millions of people turned to digital platforms to express and process their experiences (Klug et al., 2023; Risi et al., 2020). The data these platforms produced are, in principle, an extraordinary resource for social science. Yet a platform's data infrastructure is not a neutral window onto social behaviour. It is shaped by algorithmic processes, design choices, and demographic biases that researchers must understand before any substantive analysis can proceed. This thesis takes that challenge seriously. Its starting point is an evaluation of what TikTok's Research API can and cannot reliably tell us, before moving to the substantive social phenomenon the data is used to study: comedic content during a period of collective crisis. Humour has long served as a social resource in times of collective distress. Across cultures and historical periods, comedic expression functions not merely as entertainment but as a mechanism for negotiating shared meaning, enabling communities to articulate anxiety and reconstitute a sense of normalcy under conditions of uncertainty (Kuipers, 2002; Lockyer & Pickering, 2008). Comedy was central to the pandemic's digital response, and understanding how it operated on TikTok requires, first, that we understand the platform data through which it can be studied.

The present thesis examines TikTok as a data source and platform through which a specific social phenomenon can be studied: comedic content produced between 2020 and 2022.

TikTok is not simply another social media application. It is a qualitatively distinct communicative environment in which short-form video, algorithmic curation (Bhandari & Bimo, 2022), and participatory culture (Jenkins et al., 2016) intersect. These intersections fundamentally alter how comedic content is produced, circulated, and received. The thesis proceeds in three stages. It first evaluates whether TikTok's Research API can be trusted as a data source. It then examines how multimodal features of TikTok comedy jointly shape audience engagement, and how creator gender moderates those relationships. Finally, it investigates role-play as a specific creative strategy and the conditions under which it independently predicts engagement. Each stage depends on the one before it.

TikTok launched internationally in 2018, following ByteDance's merger with Musical.ly. It grew rapidly to become one of the most-downloaded applications globally (Zeng et al., 2021). By early 2026, the platform had accumulated approximately 1.59 to 1.9 billion global monthly active users (Slotta, 2026). Though sharing a common technological architecture, TikTok's national versions operate as distinct ecosystems shaped by divergent content moderation policies and audience norms (Zeng & Kaye, 2022). This thesis focuses on creators active in the United Kingdom and the United States between 2020 and 2022. This focus reflects the practical availability of data through TikTok's Research API for these two markets. It also reflects the value of examining comedy within two English-language contexts that differ markedly in their socio-political circumstances (Lin et al., 2023). The pandemic further sharpened these differences: the two countries diverged markedly in their public health messaging, lockdown timing, political leadership, and collective emotional register during 2020 to 2022. If national context shapes how audiences produce and receive comedic content, these divergences offer precisely the variation needed to detect those effects. Nationality is

therefore included not only as a demographic descriptor of the sample but as a variable that may moderate the relationships between content features and engagement.

What distinguishes TikTok from earlier platforms is the degree to which it integrates multiple communicative channels within a single short video (Zappavigna & Logi, 2025). Creators draw simultaneously on visual composition, spoken language, music, on-screen text, gesture, and costume to construct comedic meaning (Wang & Feng, 2023; Zappavigna & Logi, 2025). This multimodal density has a direct empirical consequence. No single feature, considered in isolation, can adequately account for how audiences respond. Yet previous research has predominantly examined individual elements such as hashtags, visual aesthetics, or audio characteristics in isolation. These unimodal approaches cannot capture synergistic effects across communicative channels (Klug et al., 2021; Schellewald, 2021; Omar & Dequan, 2020). This thesis addresses that gap by analysing how multiple feature categories jointly predict engagement across three distinct metrics: likes, comments, and shares.

The platform's recommendation algorithm bears directly on this analysis in two respects. First, TikTok's interest-graph model distributes content on the basis of predicted engagement rather than follower networks (Guinaudeau et al., 2022; Boeker & Urman, 2022). This shapes which comedic formats are viable, which are amplified, and which are effectively suppressed through reduced distribution. Second, the data available through TikTok's Research API is itself shaped by algorithmic and demographic processes that are not transparent to researchers. Paper 1 addresses this directly by subjecting the API's outputs to systematic comparative scrutiny. It establishes the conditions of validity for the data on which Papers 2 and 3 depend.

The pandemic intensified these dynamics in ways that are relevant to the present analysis. Lockdown conditions restricted creators' access to external locations, collaborators, and physical props. One might expect such constraints to have produced a narrowing of comedic

style. Paper 3 investigates whether creators responded by developing alternative strategies, most notably role-play and character embodiment, and whether such strategies are associated with differential engagement. Notably, explicitly COVID-themed content accounted for less than ten per cent of humour in the sample overall (Paper 1; Paper 2). Pandemic-related themes were more often embedded within everyday narratives than foregrounded as explicit subject matter. Crisis shaped the context of production; it did not dominate its content.

Three specific gaps in the existing literature motivate this research. First, TikTok's Research API has been widely adopted as a data source without systematic evaluation of its coverage, reliability, or demographic representativeness. Second, engagement research on TikTok has overwhelmingly focused on political, scientific, and health-related content. Comedic content differs from these genres in ways that have direct analytical implications. Comedy prioritises entertainment over information and grants creators substantially greater expressive and performative freedom. Comedic engagement is driven by emotional resonance and social bonding rather than informational utility (Olah & Ford, 2021; Zekavat, 2025). Comedy audiences share content to provoke laughter and connection, not primarily to inform or persuade. These distinctions mean that the multimodal features driving engagement in comedy content may differ considerably from those identified in political or scientific contexts. Yet comedy remains substantially understudied relative to its prevalence on the platform. Third, role-play as a comedic production strategy has received almost no quantitative attention. The degree to which creator gender moderates engagement effects remains poorly understood across both multimodal features and creative strategies. Each paper in this thesis addresses one of these gaps directly.

The thesis makes its contribution through three interrelated empirical studies. Paper 1 critically evaluates TikTok's Research API as a data source for social science research. It demonstrates systematic biases that compromise validity and advocates multi-method

auditing as a necessary corrective (Bai & Gu, 2026). Paper 2 examines how visual, audio, linguistic, and platform features jointly predict three engagement outcomes across 741 comedic videos from 14 creators (Paper 2). Paper 3 investigates role-play as a specific comedic production strategy, identifying its production signatures and estimating its independent effects on engagement (Paper 3). It also examines how creator gender moderates those effects. Together, the three studies draw on a shared dataset spanning a 36-month period that captures TikTok comedy around the pandemic.

The sections that follow review the theoretical literature and identify the gaps motivating each paper, describe the shared methodological approach, and outline each paper's research questions and aims. They also explain how the three studies form a coherent research programme.

2. Theoretical Framework

The theoretical frameworks drawn upon in this thesis serve two related purposes. They situate the research within established bodies of knowledge and help to identify what those bodies of knowledge have yet to address. This section reviews five conceptual areas: humour as social practice, multimodal engagement, performance theory, gender performativity, and platform affordances. Each subsection moves from foundational theory to empirical applications within social media research, before identifying the specific gaps each paper addresses. A synthesis of those gaps appears in Section 2.7.

2.1 Humour as Social Practice

Classical theories of humour were developed largely in relation to verbal jokes and literary comedy. These include incongruity theory, relief theory, and superiority theory (Beck & Spencer, 2025; Meyer, 2000; Morreall, 2014). Whilst they retain explanatory value in certain contexts, they share a tendency to locate humour within the text itself. They treat comedic

success as a property of content rather than of the interaction between content, audience, and context. This proves inadequate for short-form digital video, where meaning is produced across multiple simultaneous communicative channels. Audience response, measured through engagement metrics, constitutes the primary evidence of whether comedy has worked.

Treating humour as a social practice offers a more productive starting point (Kuipers, 2002; Lockyer & Pickering, 2008). This perspective attends to the interaction of creators, audiences, platforms, and cultural contexts in producing comedic meaning. Within this framework, the concept of coping humour is particularly relevant to the pandemic context of the present study. Research across multiple disciplines documents the role of humour in managing responses to stress, illness, and social disruption (Fogarty & Elliot, 2020; Malizal, 2024; Martin, 2007; Sliter et al., 2014). During COVID-19, comedy on social media provided resources for processing uncertainty and maintaining connection under conditions of isolation (Olah & Ford, 2021; Zekavat, 2025). Whether this coping function manifested in TikTok comedy content remains an open empirical question examined in Chapters 2 and 3.

The memetic perspective offers a further useful lens. Shifman (2015) applies Dawkins's (2016) concept of cultural transmission to internet content, drawing attention to how comedic formats are adopted, remixed, and adapted across creators. TikTok actively facilitates this through three participatory features. Duets allow creators to record a split-screen video alongside an existing video and respond to it directly. Stitches allow creators to clip and incorporate a segment of another creator's video as a starting point for their own. Sound remixing allows any audio from a published video to be extracted and reused as a soundtrack by other creators. These functions are, in a quite literal sense, memetic mechanisms. They replicate and transform comedic templates through collective creative engagement, and their use as platform affordances forms part of the analytical framework of Chapters 3 and 4.

2.2 Multimodal Engagement: An Empirical Orientation

TikTok videos integrate visual composition, speech, music, gesture, on-screen text, and physical performance within a compressed temporal frame. The simultaneous operation of these channels means that no single feature, examined in isolation, can adequately account for how audiences respond. This has direct implications for how engagement is conceptualised and studied.

Most existing studies have examined content features in isolation. They focus on single modalities such as audio (Wahid et al., 2022), visual composition (Li et al., 2021), or linguistic characteristics (Cheng & Li, 2024) without considering combined effects. Xiao et al. (2025) represent a notable exception, employing multi-method analysis to examine how multimodal features in short-form sales videos predict consumer engagement. Their focus on commercial content, however, limits generalisability to entertainment genres. Their modelling approach also does not account for the hierarchical structure of videos nested within creators. The term "multimodal" is used throughout this thesis in an empirical sense. It describes the multiple communicative channels present in TikTok videos and the analytical strategy of measuring features across those channels simultaneously. It reflects a commitment to systematic, quantitative measurement rather than any specific qualitative semiotic tradition. From the audience side, scholarly recognition of engagement as a multidimensional construct has grown substantially in recent years. Chen et al. (2021) demonstrate that likes, comments, and shares function as distinct forms of audience response rather than interchangeable indicators of a single construct. Each metric reflects different cognitive and behavioural processes. Cheng and Li (2024), examining TikTok news videos, find that sentiment and second-person address affect different engagement outcomes through different mechanisms. Shahbaznezhad et al. (2021) show that content format and platform context jointly shape which engagement behaviours users enact. Tenenboim (2022) establishes that news posts attract likes, comments, and shares through qualitatively different content characteristics.

Treating these metrics as equivalent, Tenenboim argues, conceals important variation in how audiences engage with digital content. This body of work makes a strong case for modelling engagement outcomes separately, which is precisely the approach adopted in Papers 2 and 3.

2.3 Performance Theory and the Dramaturgical Framework

Goffman's (1990) dramaturgical framework holds that social life is organised through the performance of self-presentation before real or imagined audiences. Individuals construct personas, manage impressions, and navigate the boundaries between front-stage and back-stage behaviour. Applied to digital media, this framework has generated substantial empirical work on how social media users construct and manage their self-presentation online (Barta & Andalibi, 2021; Duffy et al., 2021; Gran, 2025). TikTok creators deploy costumes, settings, and scripted interactions to produce desired audience effects. The platform's public-facing architecture makes impression management a central dimension of content production.

Empirical research has demonstrated that self-presentation strategies on TikTok are shaped by platform affordances as much as by individual choice. Barta and Andalibi (2021) find that authenticity norms on TikTok create distinctive expectations for managing the front-stage and back-stage boundary. Audiences penalise content that appears overly polished or performatively constructed. Gran (2025) argues that short-form creators adopt a mode of "performing not-not-me," a mode in which the performative frame is explicit and acknowledged whilst the creator's authentic identity remains legible beneath it. These studies suggest that Goffman's framework requires extension when the performance itself is foregrounded as part of the creative product.

This extension is precisely what role-play comedy demands. In role-play, creators explicitly embody characters distinct from their everyday personas, making the gap between performer and character a deliberate source of comedic meaning. This differs fundamentally from the impression management Goffman originally theorised. The question of whether explicit

character performance shapes engagement differently from other comedic content has not been systematically addressed in the literature. Paper 3 addresses this question directly.

2.4 Gender Performativity

Butler's (1988) conception of gender performativity holds that gender is constituted through repeated, citational acts rather than expressing a pre-existing identity. Gender is not a fixed property of persons but is enacted through habitual acts of behaviour, expression, and performance, each of which cites and reproduces established social norms. Departures from normative gender performance attract social attention precisely because they deviate from established expectations. Where Goffman treats gender as one attribute among many that performers manage, Butler foregrounds its compulsory and iterative nature as a structuring social force.

The distinction between performing gender in everyday life and performing gender in video content is non-trivial. In Goffman's account, gender is performed through everyday interaction and managed within specific situational contexts. In TikTok comedy, it is a deliberate creative choice, produced for a public audience and subject to algorithmic amplification. On such a platform, gender-norm transgression is not merely socially visible. It may also be amplified or suppressed through differential algorithmic engagement. The relationship between gender performance and digital audience response is therefore both a theoretical and an empirical question.

Empirical research on gender and social media engagement has begun to address this, though significant gaps remain. Thelwall and Foster (2021) find that gender-polarised content on YouTube receives fewer views on average. This suggests that gender-norm transgression can carry engagement costs as well as benefits. Fosch-Villaronga et al. (2021) demonstrate that gender inferences made by platforms and audiences alike shape the visibility and reception of social media content. Wu et al. (2025) show that CEO gender and content orientation interact

in shaping engagement on microblogs, with gender-incongruent styles attracting heightened attention. Research has also shown that male and female creators differ systematically in their multimodal production strategies, employing distinct combinations of visual, audio, and linguistic features (Zhang et al., 2025). These production differences provide the theoretical basis for expecting gender to moderate the relationship between content features and engagement. The extent to which this moderation operates in TikTok comedy specifically has not been systematically examined. Nor has the question of whether male and female creators derive differential benefits from role-play.

These gaps motivate two specific analytical questions in this thesis. Paper 2 asks whether and how creator gender moderates the relationship between multimodal features and three distinct engagement outcomes. Paper 3 asks whether the engagement effects of role-play differ between same-gender and cross-gender role-play. It also asks whether those effects are further moderated by creator gender. Together, these questions treat gender not as a demographic control variable but as a theoretically motivated structural moderator. Its effects are expected to differ across engagement metrics, content types, and performance strategies.

Gender stereotypes are central to understanding why creator gender should function as a moderator rather than a simple demographic covariate. Stereotypes establish culturally shared expectations about how men and women perform, speak, and carry themselves (Charlesworth et al., 2021; Eagly & Karau, 2002; Morgenroth & Ryan, 2021). When a creator's content aligns with those expectations, audiences process it fluently, with few friction points. When content departs from them, something more complex occurs. Departures can attract heightened attention and generate novel forms of engagement, but they can equally produce resistance, discomfort, or reduced amplification. The direction of the effect depends on the nature of the departure and the social licence available for it within a given context. This is especially pertinent to gender role-play, in which the departure from normative expectation is

deliberate, visible, and genre-embedded. Comedy is one context in which that social licence is unusually broad. The humorous frame explicitly signals that what follows may invert, exaggerate, or transgress normative expectations. Audiences of comedy not only permit such transgressions but actively anticipate them as a source of comedic pleasure (Kuipers, 2002). This matters for gender performance in particular.

The mechanism through which comedy enables such transgression can be understood through two theoretical accounts. Butler (1988) argues that gender parody has subversive potential by making visible the performative and constructed nature of gender identity. Comedic character performance operates as precisely this kind of parody. By foregrounding the gap between performer and persona, it renders visible what normative gender performance ordinarily conceals. Comedy as a social practice also carries an institutionalised licence for norm transgression that other communicative contexts do not (Kuipers, 2002). Character embodiment within comedy therefore allows creators to enact gender-transgressive identities that everyday interaction would not permit. Audiences decode such enactments as comedic rather than violating, guided by the generic expectations that the humorous frame establishes (Morgenroth & Ryan, 2021). The comedic frame thus functions as a structural enabler of gender norm departure, not merely a coincidental context in which such departures happen to occur.

A male creator adopting a stereotypically feminine voice or character uses the comedic frame to render the transgression legible as entertainment. So does a female creator enacting hyper-masculine aggression. Whether audiences reward or penalise such departures is an empirical question that prior research has not resolved. Whether the effect differs between male and female creators enacting the same strategy is equally unresolved. This reasoning motivates the inclusion of creator gender as a structural moderator in both Papers 2 and 3. Gender is not included as a control variable to be partialled out. It is an analytically meaningful dimension

along which the relationship between content features and audience response is expected to vary.

2.5 Platform Affordances and Algorithmic Mediation

The final theoretical strand concerns the role of platforms as active participants in shaping cultural production. Wellman et al. (2006) introduced the concept of networked affordances, later applied to digital media by boyd (2010), Marwick and boyd (2014), and Bucher and Helmond (2018). This tradition treats platform features not as neutral tools but as constraints and enablements that shape what is possible, legible, and rewarded. TikTok's video duration limits, duet and stitch functions, sound library, and algorithmic recommendation system all bear on what comedic content is viable and what kinds of engagement it generates.

TikTok's interest-graph model is particularly consequential. By prioritising engagement over social connections, the algorithm amplifies certain content characteristics and suppresses others (Guinaudeau et al., 2022; Boeker & Urman, 2022). Engagement metrics are therefore not simply reflections of audience preference. They are shaped by algorithmic amplification patterns that researchers cannot directly observe. The data available through TikTok's Research API is itself a product of these same processes. The validity of any research programme relying on that data depends on a prior understanding of what the API can and cannot reliably represent. This is the concern that motivates Paper 1.

Platform data infrastructure has attracted growing scholarly attention, driven by recognition that platform-provided datasets are shaped by design choices, demographic biases, and commercial interests (Boeker & Urman, 2022; Guinaudeau et al., 2022). The TikTok Research API has been adopted across a growing body of studies without systematic evaluation of its coverage, consistency, or demographic representativeness. Findings produced from biased or incomplete data carry those biases forward into theoretical and applied conclusions.

Addressing this gap is not merely a procedural concern. It is a condition of validity for evidence-based research on TikTok.

2.6 Integrating the Frameworks

These five conceptual threads converge on a shared set of concerns about how comedic content operates within algorithmically mediated digital environments. Each framework draws attention to the interaction between communicative content and the conditions under which it is produced and received. Humour theory examines how comedic meaning is produced through the interaction of creators and audiences in specific cultural contexts. Multimodal engagement research asks which combinations of communicative features drive distinct audience behaviours. Performance theory examines how creative strategies of self-presentation and character embodiment shape what audiences perceive as authentic or compelling. Gender performativity theory asks how gender norms are cited and transgressed through creative performance, and what consequences those acts carry in algorithmically mediated spaces. Platform affordance scholarship asks how technical and algorithmic architecture shapes what is possible and what becomes visible.

No single framework is treated as definitive, and no single paper deploys all five simultaneously. They are held in productive relation, each informing a different dimension of the research questions addressed across the three papers.

2.7 Gaps in the Literature

The foregoing review identifies three substantive gaps in existing research, each motivating a distinct paper in this thesis.

The first gap concerns the reliability and representativeness of TikTok's Research API as a data source for social science research. Despite its growing adoption, the API has not been subjected to systematic comparative evaluation against alternative data collection methods. Its coverage, demographic skews, and internal consistency remain unknown, yet these properties

bear directly on the validity of any findings derived from it. Paper 1 addresses this gap by conducting the first systematic audit of the API's outputs against a commercial analytics service and manual data collection.

The second gap concerns the multimodal predictors of engagement in TikTok comedy content. No published study has modelled the joint predictive effects of multiple feature categories on multiple engagement outcomes simultaneously within a comedy-specific dataset. This fragmented approach cannot capture synergistic or compensatory relationships between modalities. Nor can it reveal how the same features produce different effects across distinct engagement behaviours. Paper 2 addresses this gap by modelling features across five communicative dimensions simultaneously, with separate models for likes, comments, and shares.

The third gap concerns role-play as a comedic production strategy and the moderation of engagement effects by creator gender. Role-play accounts for a substantial proportion of TikTok comedy content, yet its production signatures and independent engagement effects have not been quantitatively examined. Whilst gender has been recognised as relevant to social media engagement, its role as a structural moderator of multimodal feature effects remains poorly understood. Its interaction with creative strategies such as role-play is equally underexamined. Paper 3 addresses both aspects of this gap through a two-stage analysis. It first identifies the production characteristics of role-play content, then estimates its independent engagement effects, with creator gender modelled as a systematic moderator throughout.

3. Methodological Positioning

The three papers share a common dataset and a broadly quantitative orientation. Yet each deploys a distinct analytical strategy suited to its particular research question. Together, they

constitute a coherent methodological progression. Paper 1 evaluates the reliability of the data infrastructure. Paper 2 models multimodal predictors of engagement across the full corpus. Paper 3 isolates the independent effects of a specific creative strategy within that same dataset. This section describes the methodological rationale underpinning each stage, with emphasis on framework alignment rather than procedural detail.

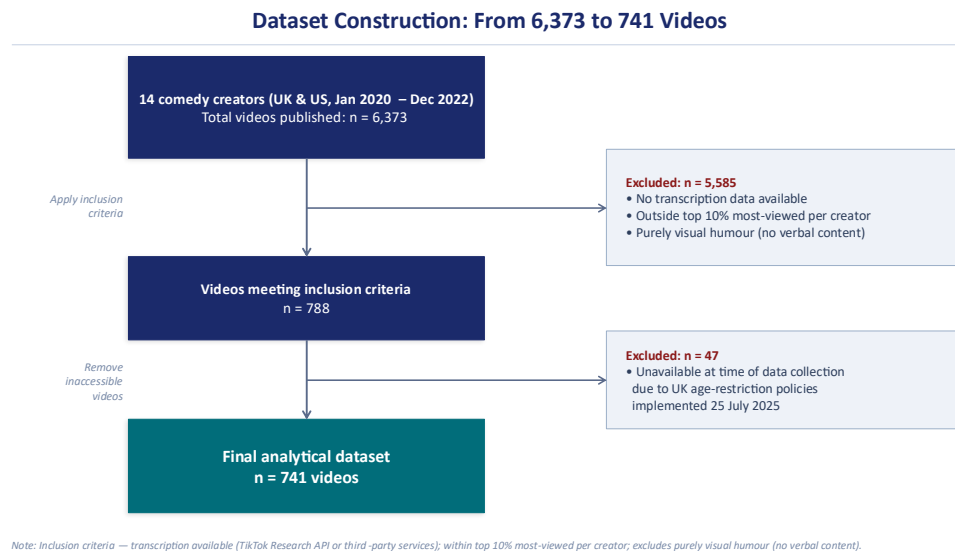
3.1 Computational Social Science and Platform Data

The dataset at the heart of all three papers was constructed through a deliberate multi-method process. Three complementary sources were combined: the official TikTok Research API, a commercial third-party analytics service (Analisa.io), and systematic manual verification. This triangulated approach was not simply a contingency. It was a principled response to documented instability in platform-provided data (Bai & Gu, 2026). Relying on any single source would have introduced unexamined biases into the subsequent analyses.

The corpus comprises 6,373 videos published by 14 comedy creators active on TikTok between January 2020 and December 2022. Creators were selected purposively, using Google search visibility and viral content indicators to identify figures sufficiently prominent to engage with broader cultural narratives. The sample balances nationality and gender, with six UK-based and six US-based creators, each group comprising three male and three female accounts. The focus on UK and US creators reflects the availability of Research API data for these two markets. It also reflects the analytical value of examining comedy within two English-language contexts that differ markedly in their socio-political circumstances (Lin et al., 2023). Sampling drew on established approaches to purposive selection in social media research (Bhandari & Bimo, 2022). Two accounts were treated as optional owing to classification ambiguities or incomplete temporal coverage. Videos were included on two conditions: the availability of transcription data and membership of each creator's top ten per cent most-viewed content. Videos consisting entirely of non-verbal humour were excluded.

After removing 47 videos inaccessible due to UK age-restriction policies, the final dataset comprised 741 videos with complete multimodal feature extraction (Figure 1).

Figure 1 Data Construction



Paper 1 subjects this data infrastructure to critical scrutiny (Bai & Gu, 2026). Drawing on computational social science frameworks, it evaluates the API’s reliability, demographic representativeness, and internal consistency relative to alternative collection methods. The findings establish the evidential conditions under which the data used in Papers 2 and 3 can be trusted. This grounding is not incidental to the thesis. It is a prerequisite for any valid inference drawn from the shared corpus.

3.2 Multimodal Feature Extraction and Modelling

Paper 2 investigates how features across multiple communicative dimensions jointly predict three engagement outcomes: likes, comments, and shares. Answering this question required a systematic feature extraction procedure and a modelling approach capable of handling non-linearity, hierarchical data structure, and gender moderation simultaneously.

Feature extraction proceeded across five dimensions. Visual features, including frame brightness, contrast, and rule-of-thirds compositional alignment, were extracted using R-based image analysis tools (Barthelme, 2015; Ooms, 2016). Audio features covered root mean

square amplitude (average sound energy), zero-crossing rate (audio texture and rhythmic variation), speech rate (syllables per second), and pitch (fundamental vocal frequency). Raw pitch values were converted from Hertz to semitones following established practice in prosody research (Shi et al., 2023). Linguistic features included type-token ratio (lexical diversity), BNC word frequency (lexical accessibility), sentiment scores, pronoun ratios, and mean length of utterance. Platform affordances, comprising filter usage and duet or stitch engagement, were hand-coded by direct observation, as automated extraction cannot reliably capture these platform-specific practices. Creator gender and nationality were included as demographic covariates throughout.

The analytical framework draws on Generalised Additive Mixed Models (GAMMs) with creator-specific random intercepts, estimated using restricted maximum likelihood (Wood & Scheipl, 2009). GAMMs were selected for two interconnected reasons. First, they accommodate non-linear relationships between content features and engagement without requiring the functional form to be pre-specified. Second, they address the nested structure of videos within creators through random intercepts, a feature that prior research has routinely neglected. Separate models were fitted for likes, comments, and shares, reflecting the theoretical expectation that these metrics represent qualitatively distinct audience behaviours rather than a single engagement construct.

3.3 Studying Performance and Production Elements

Paper 3 narrows the analytical focus to role-play, examining a comedic strategy in which creators explicitly embody characters distinct from their everyday personas. The central methodological challenge is disentanglement. Role-play content tends to employ more elaborate production elements, and any observed engagement effect could therefore reflect those correlated characteristics rather than the performance strategy itself. Paper 3 addresses

this by measuring production elements comprehensively and controlling for them statistically before estimating the independent contribution of role-play.

Role-play was coded as a three-level categorical variable: no role-play, same-gender role-play, and cross-gender role-play. Same-gender role-play involves portraying a character of the creator's own gender; cross-gender role-play involves portraying a character of a different gender. This distinction matters both theoretically and empirically, as the two types differ in their production characteristics and engagement profiles. Production elements were hand-coded by direct observation, covering costume design, accessories, hairstyling, prosthetic makeup, green screen usage, setting changes, and accent variation.

The analytical approach was tailored to a two-stage design, drawing on Linear Mixed-Effects Models and GAMMs to address two connected research questions. The first stage identifies which production elements distinguish role-play from non-role-play content, with False Discovery Rate correction applied to control for multiple comparisons. The second stage estimates whether role-play independently predicts engagement above and beyond its associated production characteristics. Creator gender is modelled as a systematic moderator throughout both stages, testing whether male and female creators experience differential production signatures and differential engagement effects from the same creative strategy.

This staged approach, from infrastructure audit to multimodal prediction to focused performance analysis, reflects the cumulative logic of the thesis as a whole. Each paper builds on the methodological foundation established by its predecessor. The shared dataset enables the findings of each paper to speak to one another directly, in ways that separately collected corpora could not support.

4. Overview of the Three Papers

TikTok transformed how millions of people produced and consumed comedy during the COVID-19 pandemic. By 2022, it had consolidated its position as one of the most culturally significant communicative environments on the planet. Comedic content was among its most

distinctive and widely shared cultural forms, and it is that content, its structure, its reception, and its underlying data, that this thesis investigates.

Chapters 3, 4 and 5 comprising this thesis address the overarching research programme from three complementary angles: data infrastructure, multimodal engagement, and creative performance. Each is self-contained and publishable as a standalone journal article. Yet the three are bound together by a shared dataset, a shared context, and a shared commitment to treating engagement as a multidimensional rather than a unitary phenomenon. What follows sets out the motivating question, research questions, and intended contribution of each paper.

4.1 Paper 1: Harnessing Big Data, Hindered by Bias: Evaluating TikTok Research API for Fair and Optimal Social Sciences

The first paper addresses a question that is prior to any substantive analysis of TikTok content.

Can the platform's own Research API be trusted as a data source for social science research?

The question is more pressing than it might initially appear. As digital platforms increasingly serve as primary archives for studying societal trends, the biases embedded in platform-provided data propagate directly into the research itself. A data source that systematically under-represents certain creators or misreports engagement metrics shapes the cumulative knowledge base of an entire field.

Paper 1 addresses this concern by subjecting the TikTok Research API to the first systematic comparative evaluation of its kind. Its central research question is as follows.

RQ: To what extent does TikTok's Research API provide reliable, demographically representative, and internally consistent data for social science research, and how do its outputs compare with those obtained through a commercial third-party analytics service and manual data collection?

The evaluation examines four specific dimensions: transcription coverage, demographic representativeness, engagement metric consistency across sources, and sensitivity of the resulting dataset to variations in inclusion criteria.

The paper's contribution is primarily methodological, but its implications extend considerably beyond methodology. By documenting the conditions under which API-derived data can and cannot be trusted to represent platform content and engagement patterns, it establishes the evidential foundation for Papers 2 and 3. It advocates for integrated auditing protocols combining API data with third-party verification and manual calibration as standard practice in platform-based social science.

4.2 Paper 2: Character Performance and Engagement: Multimodal Production Elements and Gender Effects in TikTok Role-Play Comedy

The second paper addresses a gap at the intersection of multimodal communication research and the study of social media engagement. Existing studies have overwhelmingly examined individual content features in isolation, producing fragmented findings that cannot account for synergistic relationships between communicative channels. At the same time, most studies treat likes, comments, and shares as interchangeable proxies for a single engagement construct. This assumption obscures important variation in how different features drive different audience behaviours. Paper 2 addresses both of these limitations through two research questions.

RQ1: How do visual composition, audio context, linguistic markers, and platform affordances jointly predict like count, comment count, and share count in TikTok comedy videos?

RQ2: To what extent does creator gender moderate the relationships between multimodal content features and each of the three engagement outcomes?

The paper's contribution lies in demonstrating whether and how multimodal feature configurations predict distinct engagement outcomes. The analytical design challenges the

routine practice of collapsing likes, comments, and shares into a single outcome measure, and tests whether creator gender moderates multimodal feature effects, a question that prior single-metric research has been unable to address. Together, the findings advance multimodal communication theory whilst providing practical guidance for content creators navigating algorithm-mediated platforms.

4.3 Paper 3: Character Performance and Engagement: Multimodal Production Elements and Gender Effects in TikTok Role-Play Comedy

Gender Effects in TikTok Role-Play Comedy

The third paper narrows the analytical focus to role-play, examining a comedic strategy in which creators explicitly embody characters distinct from their everyday personas. Role-play is prevalent in the dataset, accounting for 57.4 per cent of videos. Yet prior scholarship on its effects has been almost entirely qualitative or descriptive. No published quantitative study has examined which production elements characterise role-play content, whether role-play independently predicts engagement, or whether its effects differ between same-gender and cross-gender variants. Paper 3 addresses these gaps through two interconnected research questions.

RQ1: Which production elements systematically distinguish role-play from non-role-play content in TikTok comedy videos, and do these patterns differ between same-gender and cross-gender role-play?

RQ2: Does role-play independently predict audience engagement above and beyond its associated production elements, and are its effects on likes, comments, and shares moderated by role-play type and creator gender?

The paper makes two connected contributions. Empirically, it provides the first quantitative account of role-play's production signatures and independent engagement effects in TikTok comedy. Theoretically, it extends Goffman's (1990) dramaturgical framework into short-form

digital performance contexts, arguing that explicit character embodiment constitutes a distinct mode of performance. This mode differs fundamentally from the impression management Goffman originally theorised, and its consequences for engagement have not previously been examined at scale.

The three papers are not simply related by their shared dataset. They are connected by a cumulative analytical logic. Paper 1 establishes the evidential foundation by auditing the data infrastructure. Paper 2 builds on that foundation by identifying the general multimodal architecture of engagement across the full corpus. Paper 3 then uses that architecture as a control structure within which to isolate the effects of one specific creative strategy. The sequence is progressive by design, and the shared dataset ensures that the findings of each paper speak directly to those of the others.

5. Research Aims and Scope

TikTok comedy during COVID-19 is not a peripheral cultural phenomenon. It attracted hundreds of millions of viewers, generated creative strategies adapted to the specific constraints and affordances of a novel platform, and raised substantive questions about how digital comedy works, who it works for, and how it can be reliably studied. The overarching aims of this thesis are situated within that context.

Chapters 3, 4 and 5 in this thesis are connected by a shared dataset, a shared temporal context, and a shared set of concerns about how communicative features, platform infrastructure, and creative strategy interact in shaping audience engagement with comedic content on TikTok. Taken individually, each paper addresses a discrete and bounded research problem. Taken together, they constitute a programme of inquiry that moves from the foundational question of data reliability, through the modelling of multimodal engagement, to the analysis of a specific creative strategy within that multimodal framework.

The overarching aims of the thesis can be stated as follows. First, to critically evaluate the data infrastructure on which platform-based social science depends, and to establish the conditions under which TikTok Research API data can be used with confidence. Second, to determine how features spanning multiple communicative dimensions jointly predict three distinct engagement outcomes in TikTok comedy content, and to examine the extent to which those relationships are moderated by creator gender. Third, to investigate role-play as a specific comedic production strategy, identifying its characteristic production signatures and estimating its independent effects on engagement across same-gender and cross-gender variants, with creator gender again modelled as a structural moderator.

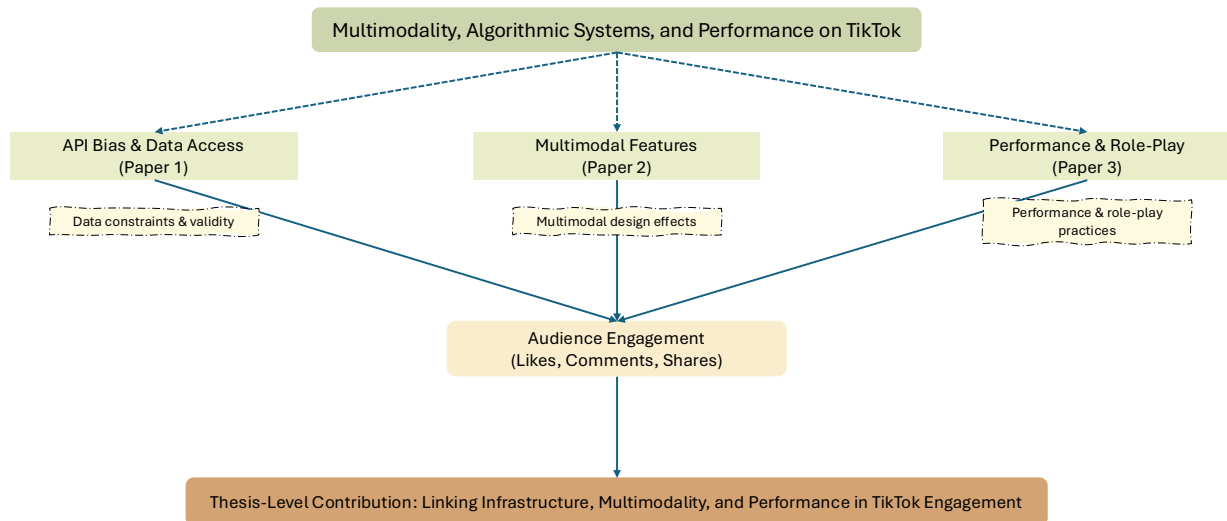
These aims speak to a set of broader questions that are likely to remain important well beyond the pandemic period that frames the data. How should platforms be held accountable for the data they make available to researchers? How do multiple communicative channels interact in shaping the reception of short-form video content? And how do creative strategies, enacted within and constrained by algorithmic environments, succeed or fail in generating different kinds of audience engagement? This thesis does not claim to resolve all of these questions. It is designed, rather, to advance each of them through evidence that is systematic in its construction and attentive to the complexity they demand.

6. Thesis Structure

This thesis is submitted in journal-article format, comprising three published or publishable papers accompanied by the present Introduction and a concluding chapter. Figure 2 illustrates the overall structure and the relationships between its components. The Introduction, the present chapter, has established the intellectual context and motivation for the research, outlined the theoretical frameworks that inform interpretation of the findings, described the methodological approach shared across the three papers, provided an overview of each

paper's question and principal findings, and articulated the integrative contribution of the thesis as a whole.

Figure 2 Thesis Structure



The Conclusion synthesises the findings across all three papers, evaluates the extent to which the three research questions have been answered, reflects on the limitations of the thesis as a whole, and identifies directions for future research into digital comedy, platform-based engagement, and the methodological challenges of studying culturally significant data at scale.

References (Introduction):

- Bai, D., & Gu, Y. (2026). Harnessing Big Data, Hindered by Bias: Evaluating TikTok Research API for Fair and Optimal Social Sciences. *Social Science Computer Review*, 08944393251413277. <https://doi.org/10.1177/08944393251413277>
- Barthelme, S. (2015). *Imager: Image processing library based on 'cimg'* [Data set]. <https://doi.org/10.32614/CRAN.package.imager>
- Beck, D., & Spencer, A. (2025). (Un)Funny against all odds: The changing landscape of humour in politics. *Alternatives: Global, Local, Political*, 50(1), 3–17. <https://doi.org/10.1177/03043754241290911>
- Bhandari, A., & Bimo, S. (2022). Why's everyone on tiktok now? The algorithmized self and the future of self-making on social media. *Social Media + Society*, 8(1), 20563051221086241. <https://doi.org/10.1177/20563051221086241>
- Boeker, M., & Urman, A. (2022). An Empirical Investigation of Personalization Factors on TikTok. *Proceedings of the ACM Web Conference 2022*, 2298–2309. <https://doi.org/10.1145/3485447.3512102>
- boyd, danah. (2010). Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In *A networked self: Identity, community and culture on social network sites* (1st Edition). Routledge.
- Bucher, T., & Helmond, A. (2018). The Affordances of Social Media Platforms. In *The Affordances of Social Media Platforms*. Sage Publications. <https://dare.uva.nl/search?identifier=149a9089-49a4-454c-b935-a6ea7f2d8986>
- Butler, J. (1988). Performative acts and gender constitution: An essay in phenomenology and feminist theory. *Theatre Journal*, 40(4), 519. <https://doi.org/10.2307/3207893>
- Charlesworth, T. E. S., Yang, V., Mann, T. C., Kurdi, B., & Banaji, M. R. (2021). Gender stereotypes in natural language: Word embeddings show robust consistency across

- child and adult language corpora of more than 65 million words. *Psychological Science*, 32(2), 218–240. <https://doi.org/10.1177/0956797620963619>
- Chen, Q., Min, C., Zhang, W., Ma, X., & Evans, R. (2021). Factors driving citizen engagement with government tiktok accounts during the covid-19 pandemic: Model development and analysis. *Journal of Medical Internet Research*, 23(2), e21463. <https://doi.org/10.2196/21463>
- Cheng, Z., & Li, Y. (2024). Like, comment, and share on tiktok: Exploring the effect of sentiment and second-person view on the user engagement with tiktok news videos. *Social Science Computer Review*, 42(1), 201–223. <https://doi.org/10.1177/08944393231178603>
- Dawkins, R. (2016). *The selfish gene* (40th anniversary edition). Oxford University Press.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. <https://doi.org/10.1037/0033-295X.109.3.573>
- Fogarty, M., & Elliot, D. L. (2020). The role of humour in the social care professions: An exploratory study. *The British Journal of Social Work*, 50(3), 778–796. <https://doi.org/10.1093/bjsw/bcz027>
- Goffman, E. (1990). *The presentation of self in everyday life* (1. Anchor Books ed., rev. ed). Anchor Books.
- Guinaudeau, B., Munger, K., & Votta, F. (2022). Fifteen Seconds of Fame: TikTok and the Supply Side of Social Video. *Computational Communication Research*, 4(2), 463–485. <https://doi.org/10.5117/CCR2022.2.004.GUIN>
- Jenkins, H., Itō, M., & boyd, danah. (2016). Participatory culture in a networked era: A conversation on youth, learning, commerce, and politics. Polity Press.
- Klug, D., Evans, M., & Kaufman, G. (2023). How TikTok served as a platform for young people to share and cope with lived COVID-19 experiences. *MedieKultur: Journal of*

- Media and Communication Research*, 38(73), 152–170.
<https://doi.org/10.7146/mk.v38i73.128463>
- Klug, D., Qin, Y., Evans, M., & Kaufman, G. (2021). Trick and Please. A Mixed-Method Study On User Assumptions About the TikTok Algorithm. *13th ACM Web Science Conference 2021*, 84–92. <https://doi.org/10.1145/3447535.3462512>
- Kuipers, G. (2002). Media culture and Internet disaster jokes: Bin Laden and the attack on the World Trade Center. *European Journal of Cultural Studies*, 5(4), 450–470.
<https://doi.org/10.1177/1364942002005004296>
- Lin, J., Swart, J., & Zeng, G. (2023). Theorising TikTok cultures: Neuro-images in the era of short videos. *Media, Culture & Society*, 45(8), 1550–1567.
<https://doi.org/10.1177/01634437231202167>
- Lockyer, S., & Pickering, M. (2008). You Must Be Joking: The Sociological Critique of Humour and Comic Media: Humour and Comic Media. *Sociology Compass*, 2(3), 808–820. <https://doi.org/10.1111/j.1751-9020.2008.00108.x>
- Lyon, D., & Coleman, R. (2023). Rupture, repetition, and new rhythms for pandemic times: Mass Observation, everyday life, and COVID-19. *History of the Human Sciences*, 36(2), 26–48. <https://doi.org/10.1177/09526951221133983>
- Malizal, Z. Z. (2024). The role of humor in stress management and psychological well-being: Coping strategies in high-stress populations. *Sinergi International Journal of Psychology*, 2(2), 80–93. <https://doi.org/10.61194/psychology.v2i2.522>
- Martin, R. A. (2007). *The psychology of humor: An integrative approach*. Elsevier Academic Press.
- Marwick, A. E., & boyd, danah. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, 16(7), 1051–1067.
<https://doi.org/10.1177/1461444814543995>

- Meyer, J. C. (2000). Humor as a double-edged sword: Four functions of humor in communication. *Communication Theory*, 10(3), 310–331.
<https://doi.org/10.1111/j.1468-2885.2000.tb00194.x>
- Morgenroth, T., & Ryan, M. K. (2021). The effects of gender trouble: An integrative theoretical framework of the perpetuation and disruption of the gender/sex binary. *Perspectives on Psychological Science*, 16(6), 1113–1142.
<https://doi.org/10.1177/1745691620902442>
- Morreall, J. (2014). Humor, philosophy and education. *Educational Philosophy and Theory*, 46(2), 120–131. <https://doi.org/10.1080/00131857.2012.721735>
- Olah, A. R., & Ford, T. E. (2021). Humor styles predict emotional and behavioral responses to COVID-19. *HUMOR*, 34(2), 177–199. <https://doi.org/10.1515/humor-2021-0009>
- Omar, B., & Dequan, W. (2020). Watch, Share or Create: The Influence of Personality Traits and User Motivation on TikTok Mobile Video Usage. *International Journal of Interactive Mobile Technologies (iJIM)*, 14(04), 121.
<https://doi.org/10.3991/ijim.v14i04.12429>
- Ooms, J. (2016). *Magick: Advanced graphics and image-processing in r* [Data set].
<https://doi.org/10.32614/CRAN.package.magick>
- Risi, E., Pronzato, R., & Di Fraia, G. (2020). Living and working confined at home: Boundaries and platforms during the lockdown. *Journal of Cultural Analysis and Social Change*, 5(2), 12. <https://doi.org/10.20897/jcasc/9253>
- Schellewald, A. (2021). Communicative Forms on TikTok: Perspectives From Digital Ethnography. *International Journal of Communication*, 15(0), 21.
<https://ijoc.org/index.php/ijoc/article/view/16414>

- Shi, J., Gu, Y., & Vigliocco, G. (2023). Prosodic modulations in child-directed language and their impact on word learning. *Developmental Science*, 26(4), e13357.
<https://doi.org/10.1111/desc.13357>
- Shifman, L. (2015). *Memes in digital culture*. CRC Press.
- Sliter, M., Kale, A., & Yuan, Z. (2014). Is humor the best medicine? The buffering effect of coping humor on traumatic stressors in firefighters. *Journal of Organizational Behavior*, 35(2), 257–272. <https://doi.org/10.1002/job.1868>
- Slotta, D. (2026, January 13). *TikTok—Statistics & facts*. statista. Statista.
<https://www.statista.com/topics/6077/tiktok/#topicOverview>
- Wang, Y., & Feng, D. (William). (2023). History, modernity, and city branding in China: A multimodal critical discourse analysis of Xi'an's promotional videos on social media. *Social Semiotics*, 33(2), 402–425. <https://doi.org/10.1080/10350330.2020.1870405>
- Wellman, B., Quan-Haase, A., Boase, J., Chen, W., Hampton, K., Díaz, I., & Miyata, K. (2006). The Social Affordances of the Internet for Networked Individualism. *Journal of Computer-Mediated Communication*, 8(3), 0–0. <https://doi.org/10.1111/j.1083-6101.2003.tb00216.x>
- Wood, S., & Scheipl, F. (2009). *Gamm4: Generalized additive mixed models using 'mgcv' and 'lme4'* [Data set]. <https://doi.org/10.32614/CRAN.package.gamm4>
- Xiao, Q., Huang, W., Qu, L., & Li, X. (2025). The impact of multimodal information features of short sales videos on consumer engagement behavior: A multi-method approach. *Journal of Retailing and Consumer Services*, 82, 104136.
<https://doi.org/10.1016/j.jretconser.2024.104136>
- Zappavigna, M., & Logi, L. (2025). How emoji make meaning and enact ambient affiliation: A social semiotic account of emoji-text relations in TikTok comments. *Social Semiotics*, 35(4), 617–637. <https://doi.org/10.1080/10350330.2024.2389520>

- Zekavat, M. (2025). Laughter in times of distress: Pandemic humor and satire in Iran. *European Journal of Cultural Studies*, 28(6), 1799–1823.
<https://doi.org/10.1177/13675494251331211>
- Zeng, J., & Kaye, D. B. V. (2022). From content moderation to *visibility moderation*: A case study of platform governance on TikTok. *Policy & Internet*, 14(1), 79–95.
<https://doi.org/10.1002/poi3.287>
- Zeng, J., Abidin, C., & Schäfer, M. S. (2021). *Research perspectives on TikTok and its legacy apps: Introduction*. <https://doi.org/10.5167/UZH-205427>
- Shi, J., Gu, Y., & Vigliocco, G. (2023). Prosodic modulations in child-directed language and their impact on word learning. *Developmental Science*, 26(4), e13357. <https://doi.org/10.1111/desc.13357>
- Shifman, L. (2015). *Memes in digital culture*. CRC Press.
- Sliter, M., Kale, A., & Yuan, Z. (2014). Is humor the best medicine? The buffering effect of coping humor on traumatic stressors in firefighters. *Journal of Organizational Behavior*, 35(2), 257–272. <https://doi.org/10.1002/job.1868>
- Slotta, D. (2026, January 13). *TikTok—Statistics & facts*. statista. Statista.
<https://www.statista.com/topics/6077/tiktok/#topicOverview>
- Wang, Y., & Feng, D. (William). (2023). History, modernity, and city branding in China: A multimodal critical discourse analysis of Xi'an's promotional videos on social media. *Social Semiotics*, 33(2), 402–425. <https://doi.org/10.1080/10350330.2020.1870405>
- Wellman, B., Quan-Haase, A., Boase, J., Chen, W., Hampton, K., Díaz, I., & Miyata, K. (2006). The Social Affordances of the Internet for Networked Individualism. *Journal of Computer-Mediated Communication*, 8(3), 0–0. <https://doi.org/10.1111/j.1083-6101.2003.tb00216.x>
- Wood, S., & Scheipl, F. (2009). *Gamm4: Generalized additive mixed models using 'mgcv' and 'lme4'* [Data set]. <https://doi.org/10.32614/CRAN.package.gamm4>

- Xiao, Q., Huang, W., Qu, L., & Li, X. (2025). The impact of multimodal information features of short sales videos on consumer engagement behavior: A multi-method approach. *Journal of Retailing and Consumer Services*, 82, 104136.
<https://doi.org/10.1016/j.jretconser.2024.104136>
- Zappavigna, M., & Logi, L. (2025). How emoji make meaning and enact ambient affiliation: A social semiotic account of emoji-text relations in TikTok comments. *Social Semiotics*, 35(4), 617–637. <https://doi.org/10.1080/10350330.2024.2389520>
- Zekavat, M. (2025). Laughter in times of distress: Pandemic humor and satire in Iran. *European Journal of Cultural Studies*, 28(6), 1799–1823.
<https://doi.org/10.1177/13675494251331211>
- Zeng, J., & Kaye, D. B. V. (2022). From content moderation to *visibility moderation*: A case study of platform governance on TikTok. *Policy & Internet*, 14(1), 79–95.
<https://doi.org/10.1002/poi3.287>
- Zeng, J., Abidin, C., & Schäfer, M. S. (2021). *Research perspectives on TikTok and its legacy apps: Introduction*. <https://doi.org/10.5167/UZH-205427>

Chapter 2: Harnessing Big Data, Hindered by Bias: Evaluating TikTok Research API for Fair and Optimal Social Sciences

Abstract

Digital platforms now serve as crucial archives for analysing societal trends, yet their research APIs pose methodological challenges. This study critically evaluates the TikTok Research API through a comparative analysis of 6,373 videos from 14 creators in the United States and the United Kingdom (2020–2022), contrasting API-derived outputs with manual collection and third-party analytics. The API demonstrated scalability, retrieving more videos than alternative methods and providing 22 variables, including eight unavailable elsewhere. However, limitations were substantial: transcriptions covered about 10% of the content, with more transcripts returned from American male creators. Engagement metrics exhibited inconsistent accuracy across data sources, with the API showing systematically lower view counts but higher comment and share counts compared to manual collection. The number of videos varied depending on sample composition, indicating that small changes in inclusion criteria could shift outcomes disproportionately. These results highlight systematic inconsistencies, underscoring the need for multi-method approaches to remain necessary despite automation. While TikTok Research API offers valuable scale and ethical compliance, its demographic biases and metadata inconsistencies compromise validity. The study advocates integrated auditing protocols and targeted API refinements to improve representativeness and accuracy in platform-based research.

Keywords: digital platform, TikTok, API, big data, data audit

Introduction

Digital platforms have become integral to contemporary social life, producing vast datasets that can be systematically examined to address pressing research questions. Social media (SM) platforms generate digital traces that offer unprecedented opportunities to study societal trends (Reisach, 2021) and human behaviour across diverse contexts. These data enable researchers to explore phenomena ranging from conflict dynamics (Zeitsoff, 2011) to polarisation in voting behaviour (Sharma et al., 2022), underscoring their relevance to computational social science (Lazer et al., 2021). The dynamic interplay of platform algorithms, user interactions, and societal events necessitates robust open science practices to ensure transparency and reproducibility (Bezjak et al., 2018; Mostar et al., 2025). Application programming interfaces (APIs) provide structured protocols for accessing these data,

facilitating large-scale analyses of networked behaviours (Acker & Kreisberg, 2020; Davidson et al., 2023).

Among social media platforms, TikTok has emerged as a focal point for researchers due to its rapid growth and appeal to younger demographics. Access to TikTok's Research API necessitates an endorsement letter, a requirement that has grown common, though the frequency of approvals remains unclear due to limited public data on application outcomes. By late 2023, TikTok expanded its Research API to European countries, offering a more permissive approach compared to platforms like Meta and X, which have curtailed academic access (Chang, 2018; TikTok, 2023). TikTok's substantial revenue growth, reaching £875 million by October 2022 (Sweney, 2022), underscores its expanding influence across cultural and political domains (Literat, 2021; Rejeb et al., 2024). The platform's rich, multimodal data, encompassing videos, audio, text, and engagement metrics, support a diverse array of research methodologies, ranging from thematic analysis (Moir, 2023; Lu & Shen, 2023) and digital ethnography (Entrena-Serrano, 2025; Haime & Biddle, 2025; Yang, 2022) to interviews (Klug et al., 2021; Liang, 2021), surveys (Chen & Zhang, 2021; Kirkpatrick & Lawrie, 2024), and machine learning applications (Agrawal, 2024; Parisi et al., 2023). However, platform-imposed restrictions arising from Terms of Service (ToS), such as data retention, pre-production, post-publication, and risks of indemnity, often constrain research validity and replicability (Bak-Coleman, 2023; Venkatagiri, 2023; related items in Appendix 1).

Additionally, significant methodological and ethical challenges continue to limit the validity of societal insights derived from TikTok research. Small sample sizes, often the result of manual scraping (e.g., Cervi & Divon, 2023), restrict representativeness. Ethical concerns also arise when data collection contravenes TikTok's Terms of Service, whether through manual or automated scraping (Thole, 2022). Unstable API performance can return data inconsistent with webpage information (Ruz et al., 2023), while reliance on third-party tools

frequently produces incomplete datasets requiring extensive cleaning and validation (Mimizuka et al., 2025). Meanwhile, despite the widespread use of API-driven data, little is known about potential algorithmic biases favouring certain demographics or about inconsistencies in metadata accuracy. This lack of understanding undermines the ability to assess the generalisability and validity of current acquisition methods, and, by extension, the robustness of findings based upon them.

Existing research highlights persistent technical and ethical challenges in using TikTok's Research API. There are three key limitations: inconsistent data delivery, rate restrictions on video retrieval, and restricted API access for independent researchers, which undermines the quality of societal insights that can be derived from platform data during critical events such as the COVID-19 pandemic. Notably, algorithmic curation and demographic disparities, including gender and cultural skews, introduce biases that call into question the fairness of trend analyses on the platform. Crucially, integrating multi-method approaches, such as combining API data with manual collection or external analytics, offers a means to address these inconsistencies. This is consistent with calls to enhance the reliability of social media research by triangulating sources, particularly when examining time-bound phenomena such as lockdown periods.

This study critically evaluates TikTok Research API by comparing API-derived data with manual collection and third-party analytics, using a dataset of 6,373 videos (after manual calibration) from 14 U.S. and U.K. creators (2020–2022). Videos were selected for their relevance to social commentary, humour, and public discourse during the COVID-19 period.

This study addresses key research questions:

RQ1. To what extent does TikTok's Research API enable effective multimodal data collection for computational social science, and how do its technical limitations and ethical constraints influence the validity of societal insights derived from such data?

RQ2. In what ways do algorithmic and demographic biases within TikTok's Research API affect the fairness of societal trend studies based on its multimodal data?

RQ3. How can integrating multi-method data collection with TikTok Research API mitigate inconsistencies to yield valid and equitable insights into societal trends?

This paper proceeds as follows: it first examines the functionalities of the TikTok Research API, then evaluates its limitations, and finally proposes methodological refinements. The literature review synthesises prior scholarship on social media APIs, covering data collection practices, the affordances and constraints of the TikTok Research API, and associated issues of data quality, governance, and ethics. The methodology outlines data acquisition from three sources, such as API retrieval, paid datasets, and manual calibration, followed by statistical evaluation of data quality across these streams. The findings present the results of the comparative analysis. The discussion then interprets these outcomes, considering the research questions, drawing out implications for platform-based data acquisition strategies. The paper concludes with recommendations to strengthen methodological rigour in future social media research.

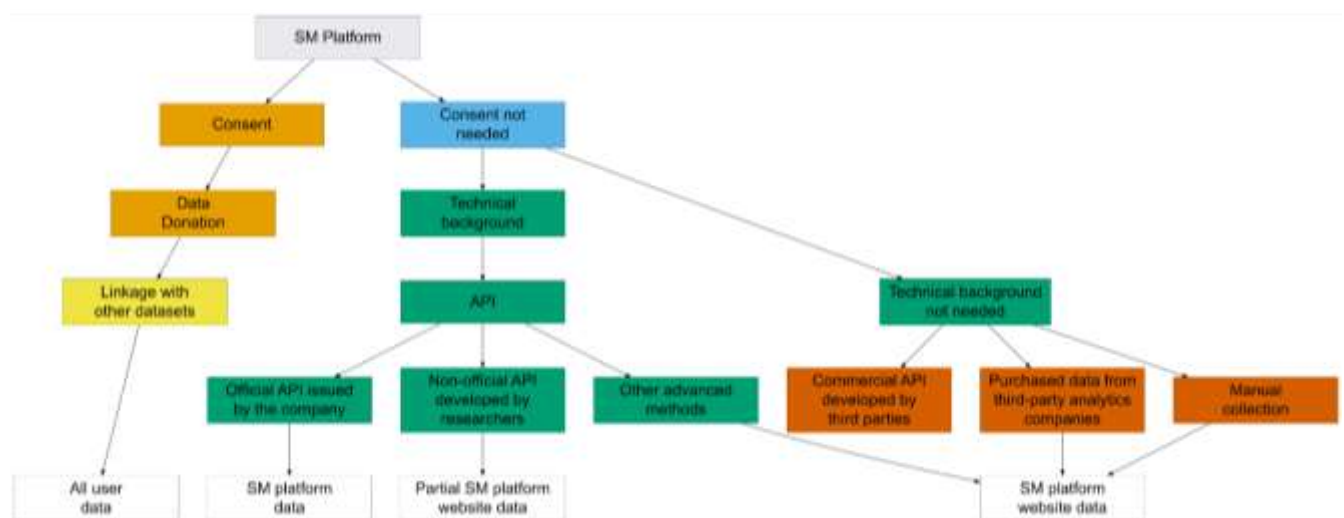
Literature Review

2.1 Data Collection in Computational Social Science

Computational social scientists typically employ one or more of four approaches to data collection: web scraping, commercial analytics services, data donation, and platform-provided APIs, with the choice shaped by participant consent and the researcher's technical expertise. Among these, APIs offer a sanctioned and legally compliant route to structured, large-scale public social media data, with clear advantages over manual collection and unofficial scraping (see Figure 1; Table 1) [Insert (Figure1) here]. Official APIs, such as those provided by TikTok, can grant access to archived content not visible on public-facing interfaces, thereby

ensuring a degree of compliance with platform terms (Davidson et al., 2023). However, restrictive usage agreements, including rate limits and potential costs, can narrow the research scope and hinder open dissemination (Ohme et al., 2024). Manual collection remains accessible to those without advanced technical skills; however, it is labour-intensive, error-prone, and poorly suited to large datasets (e.g., Cervi & Divon, 2023). Unofficial scraping and some third-party crawlers present an alternative, yet they may breach terms of service and raise significant ethical and legal concerns (Bruns, 2019). Taken together, these constraints underscore the need to critically evaluate API-based methods for fairness, validity, and their capacity to support robust social media research.

Figure 3 Accessing SM Data



Note: the diagram is built on Davidson et al. (2023)

2.2 The TikTok Research API: Promise and Pitfalls

Launched in 2023, the TikTok Research API represents a milestone in platform transparency and academic collaboration. Initial applications of the API have enabled researchers to analyse political communication, content virality, and algorithmic demographic bias through machine learning models (Corso et al., 2024; Pinto et al., 2024). The structured provision of multimodal data, including text, video, and audio metadata, marks a significant step forward compared to earlier TikTok studies reliant on ad hoc methods (Rogers & Zhang, 2024). This

has enabled analyses ranging from content engagement statistics to semiotic and linguistic interpretations (Leedham et al., 2020; Parikh, 2025).

Nevertheless, the practical implementation of the API is far from seamless. Scholars have reported incomplete and inconsistent data delivery, often receiving only a fraction of the requested materials. Burnnat and Davidson (2025) note that TikTok Research API imposes highly selective access criteria and suffers from documented data quality issues, corroborating reports of inconsistent data delivery (Corso et al., 2024; Ruz et al., 2023). Pfeffer et al. (2018) demonstrate that streaming sample APIs are neither random nor immune to distortion: deliberate flooding can manipulate topic coverage, while rate limits and batching introduce temporal artefacts and omit low-activity content. Furthermore, high-engagement posts are over-represented, challenging the notion that API data provide a faithful snapshot of platform dynamics. These constraints do not only limit the API's utility for comprehensive societal trend analysis, but also can severely impact dataset completeness, thereby compromising analytical accuracy and interpretive robustness (*ibid*). Additionally, server instability and unreliable pagination further undermine the reliability of automated data extraction processes (Mimizuka et al., 2025). If API workflows encourage fragmented or low-fidelity datasets, then the integrity of empirical findings is inevitably weakened.

Table 1 Description of each route from Figure 1, with the types of data that are typically obtained from that route

	Type of data collection	Type of data collected	Example	Pros	Cons	
	Data donation	Participants from a different dataset agree to offer their SM platform information.	Typically, all user data available on the platform (content, posts)	Researchers ask for consent from participants in a nationally representative survey of the U.K. if they can use their SM platform information (such as Twitter and LinkedIn ¹	Legal and academically ethical	A smaller sample size
	Official API	Issued by the specific platform or company as the only official way to access data	SM platform data	TikTok Research API	Legal and academically ethical	Immensely restricted to the SM platform terms of service: Specific application procedure (sometimes requires reference letters)
	Commercial API developed by third-party companies	Also known as a scraper, spider, or crawler, developed by third-party companies for commercial use	SM platform website data	TikTok Scraper by APIFY ² or TikAPI ³	Legal for the researcher and academically ethical; no technical background required	Payment (relatively affordable)
	Non-official API developed by other researchers/scientists	Developed and maintained by researchers	SM platform website data (partially)	GitHub wrapper ⁴	Free; legal for the researcher	It may not work correctly when the platform changes its algorithms to prevent scraping.
	Other advanced software or methods	Developed and maintained by experts who may not use common programming languages or software (such as R and Python)	SM platform website data	Interface (previously known as Stiftung Neue Verantwortung) ⁵	Free; legal; sometimes more data than official API	Ambiguous in research ethics ⁶
	Manual Scraping	Researchers go to the webpage and manually copy information	SM platform data	Researchers may watch a specific group of videos under a hashtag ⁷	Free, more reliable and increased familiarity with the datasets; legal; no technical background required	Labour-intensive; less effective or a smaller sample; ambiguous in research ethics
	Purchased data	Researchers can choose from diverse services in the market	SM platform data	Analisa.io	Data has already been cleaned for use with customised variables, which are legal and academically ethical; no technical background required.	Higher costs and still require minor manual calibration ⁸ .

¹ Liu, S., Sloan, L., Al Baghal, T., Williams, M., Jessop, C., & Serôdio, P. (2024). Linking survey with Twitter data: Examining associations among smartphone usage, privacy concern and Twitter linkage consent. *International Journal of Social Research Methodology*, 1–15. <https://doi.org/10.1080/13645579.2023.2299482>

² <https://apify.com/clockworks/tiktok-scraper>

³ <https://tikapi.io/>

⁴ Vombatkere, K., Mousavi, S., Zannettou, S., Roesner, F., & Gummadi, K. P. (2024). Tiktok and the art of personalization: Investigating exploration and exploitation on social media feeds. *Proceedings of the ACM Web Conference 2024*, 3789–3797. <https://doi.org/10.1145/3589334.3645600>

⁵ <https://tiktok-audit.com/about/> and <https://tiktok-audit.com/blog/2023/the-TikTok-research-API-falls-woefully-short/>

⁶ Krotov, V., Johnson, L., Murray State University, Silva, L., & University of Houston. (2020). Legality and ethics of web scraping. *Communications of the Association for Information Systems*, 47, 539–563. <https://doi.org/10.17705/1CAIS.04724>

⁷ Divon, T., & Ebbrecht-Hartmann, T. (2023). PERFORMING DEATH AND TRAUMA? PARTICIPATORY MEM(E)ORY AND THE HOLOCAUST IN TIKTOK #POVCHALLENGES. *AoIR Selected Papers of Internet Research*. <https://doi.org/20230412082800000>

⁸ Official TikTok business partners: <https://partners.tiktok.com/directory/pc/en?rid=xm39skrb37c&specialties=304%2C505%2C504%2C501%2C503%2C502%2C103%2C104>

Note: The table is built on Davidson et al. (2023)

2.3 Data Quality, Governance, and Ethics

Concerns over the epistemological trustworthiness of API-derived data extend beyond purely technical limitations. TikTok's internal moderation algorithms and content-curation systems introduce layers of bias and obfuscation rarely addressed in dataset evaluations. These 'black-box' mechanisms can distort both the authenticity and representativeness of retrieved data, creating systemic blind spots in research outputs (Corso et al., 2024).

Governance frameworks further shape who can access platform APIs, often reinforcing epistemic privilege for university-affiliated researchers. Under the European Union's Digital Services Act, applicants must demonstrate formal affiliation with a recognised academic or non-profit institution to obtain 'vetted' status for internal datasets (European Centre for Algorithmic Transparency, 2025). In practice, this requirement compels independent researchers, freelance journalists, and civil-society organisations to partner with universities, think tanks, or research consortia to meet eligibility criteria. As a result, API access remains largely confined to academic investigators, excluding many non-academic actors from direct engagement with platform data (Mimizuka et al., 2025).

Burnnat and Davidson (2025) identify critical 'audit blind-spots' in platform APIs, including TikTok's, where moderation and algorithmic processes remain opaque, an 'accountability paradox' in which platforms control both the data and the terms of its release. This lack of transparency, coupled with epistemic asymmetries favouring platform operators, raises significant ethical challenges for ensuring fairness and validity in computational social science.

A further point of contention concerns user consent and data privacy. The granularity of engagement metrics available through the API, when combined with other datasets, can enable inadvertent re-identification of individual users. Such risks create potential legal liabilities and challenge the ethical legitimacy of research reliant on these data (Mimizuka et al., 2025), especially for those who are not qualified as public figures.

In summary, while the TikTok Research API represents a notable advancement in platform-based research, the literature portrays its utility as mixed. Technical constraints, governance opacity, and ethical exclusions collectively limit its broader adoption and impact. These unresolved issues form the rationale for this study's

central aim: to conduct a robust, comparative audit of TikTok’s Research API against alternative data-collection methods, including web scraping and commercial analytics services, thereby contributing to methodological best practices in computational social science and informing future platform policy.

Methodology

This study employs a mixed-methods design to evaluate the efficacy and limitations of TikTok Research API in analysing societal trends, with a specific focus on comedic content during the COVID-19 pandemic. Data were obtained from three sources: the official TikTok API, a third-party analytics service (Analisa.io), and manual calibration, enabling a comparative assessment of accuracy, completeness, and representativeness. In doing so, the research addresses persistent gaps in platform-based studies (Davidson et al., 2023). The following subsections outline the sampling strategy, data acquisition procedures, and analytical approach, with a particular focus on methodological robustness and ethical compliance.

3.1 Sample Selection

The COVID-19 pandemic presents a critical context for examining societal responses under conditions of collective stress, revealing patterns of isolation, mental health strain, and the circulation of misinformation (Davidson et al., 2023). Comedy on TikTok, often rooted in shared social truths, provides a distinctive lens through which to interrogate cultural narratives and patterns of user engagement. As neuroimaging research indicates, the brain’s response to vividly presented content can mirror real-world experiences, harnessing neuroplasticity to capture attention and reflect the prevailing zeitgeist (Nishimoto et al., 2011; Pearson et al., 2015, 2019). Against this backdrop, prominent comedic creators were selected as a purposive sample to examine real-time societal discourse.

The raw dataset comprised 6,373 videos posted between 1 January 2020 and 31 December 2022 from 14 high-profile TikTok comedians: seven from the United Kingdom and seven from the United States, representing both genders. Creators were identified through a combination of Google search visibility and viral content indicators, ensuring coverage of influential socio-cultural narratives (see Table 2 and Appendix 2 for video counts and detailed sampling procedures, respectively). Two accounts, “@steven_he” and “@nicholas_flannery”, were designated as optional. The former could not be unambiguously classified as

British or American, identifying as Chinese Irish and residing in the United States during the study period, while the latter began uploading only in 2021, resulting in incomplete temporal coverage. Thus, the initial dataset for the following analysis includes 12 creators (6 from each country, balanced by gender in each country, totalling 5,246 videos).

Table 2 Total number of videos of selected content creators

Username handle	Gender	Nationality	No. of Videos (Research API)	No. of transcriptions (Research API)	transcription%	Period of transcriptions
lizzza	female	American	159	1	1%	2022
daniel.labelle	male	American	302	19	6%	2022
hannahstocking	female	American	606	28	5%	2022
jeremylynch	male	British	792	92	12%	2021.9~2022
adamw	male	American	632	111	18%	2021.6~2022
spencewuah	male	American	2,346	188	8%	2021.6~2022
adrianbliss	male	British	106	0	0%	N/A
steven_he	male	American (Chinese Irish actor)	199	35	18%	2021.12~2022
hayleygeorgiamorris	female	British	187	0	0%	N/A
jackjos3ph	male	British	832	0	0%	N/A
nicholas_flannery	male	British	842	13	2%	2021.2~2022
ameliadimz	female	British	141	13	9%	2021.8~2022
itscaitlinhello	female	American	160	28	18%	2021.4~2022
maddiegracejepson	female	British	580	136	23%	2022

Note: The content creators analysed in this study have not been anonymised, as they are public figures (Stevens et al., 2015). This approach parallels established conventions in SM research, where participant identities are likewise reported without pseudonymization (e.g., Alexandre et al., 2022; Arora et al., 2019).

The core analytical materials comprised video metadata (e.g., upload date, video description) (Figure2) and engagement metrics (e.g., views, likes, comments, shares). These provided both quantitative indicators and discursive content, enabling a multi-layered analysis of comedic expression and audience interaction during a period of significant social disruption.

3.2 Data Acquisition and Annotations

To enable a triangulated assessment of data fidelity, three distinct sources (Table 3) were employed: a commercial analytics platform (Analisa.io), the official TikTok Research API, and manual calibration. This multi-source design enabled systematic cross-validation of completeness, accuracy, and consistency across retrieval methods.

The first dataset was obtained in March 2023 from Analisa.io, a commercial social media analytics service (Lachief, 2023). However, access to several creator profiles, including @lizzza, @ameliadimz, @maddiegracejepson, and @hannahstocking, was subsequently lost following the service’s closure, resulting in partial discontinuity.

Table 3 Data Description from Three Data Sources

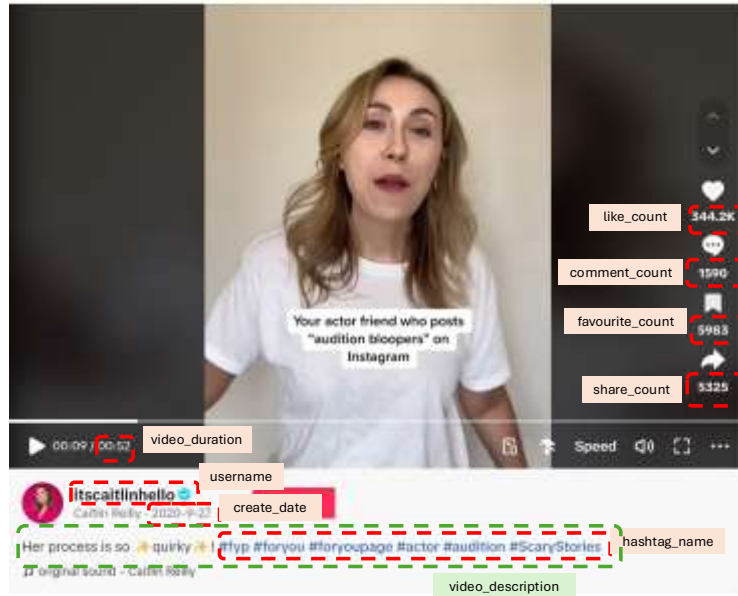
username	Analisa.io	TikTok Research API	Manual Calibration
lizzza	N/A	159	168
daniel.labelle	319	302	319
hannahstocking	N/A	606	691
jeremylynch	738	792	734
adamw	567	632	584
spencewuah	457	2,346	449
adrianbliss	69	106	113
steven_he	237	199	237
hayleygeorgiamorris	204	187	201
jackjos3ph	806	832	1,028
nicholas_flannery	902	842	890
ameliadimz	N/A	141	156
itscaitlinhello	170	160	170
maddiegracejepson	N/A	580	592
Number of Participants	10	14	14
Number of Videos	4,469	7,884	6,373
Engagement Metrics	view, likes, comments, shares	view, likes, comments, shares and transcriptions	view, likes, comments, shares
Time of collection	Mar-23	April 2024	Between April and July 2023

The second dataset was collected in April 2024 via the official TikTok Research API (the raw data are shown in Figure 3) [Insert (Figure3) here], accessed in R using packages such as httr and jsonlite for HTTP requests and JSON parsing (Pascual, 2020). Notably, fields in Figure 3 are publicly available information, which does not qualify as an “internal” dataset under the Digital Services Act (European Commission, 2025), nor do they contain sensitive or private information beyond what is already publicly accessible. All codes are available on GitHub⁹. At the time of collection, three primary endpoints were available (see the list of endpoints in Appendix 3), including video metadata, comments, and creator profiles, with video metadata forming the principal analytical focus (Figure 2) [Insert (Figure2) here]. Data were filtered by publication

⁹ https://anonymous.4open.science/r/r-tiktok-research-api-4E35/single_user_video

date (2020–2022), geographic region (United Kingdom and United States), and key content metrics (Appendix 4).

Figure 4 Sample Video by @itscaitlinhello



Note: “favourite_count” is not included in this paper due to its unavailability by the time of data collection (April 2024)

Figure 5 Example Output from API

create_time	like_count	comment_count	share_count	view_count	video_description	hashtag_names	voice_to_text	id	music_id	effect_ids	region_cod	username
28/12/2022 21:07	1340936	14533	33668	4750903	“I LOVE YOUR SON”	christmas, foryou, pro	hi. I know everyone's	7.18231E+7	7.18231E+7	0	US	itscaitlinhello
24/11/2022 18:45	68249	1114	1304	326738	Happy thanksgiving, resp	foryou, fyp, thankgavi	hey, I know it's mornin	7.16966E+7	7.16966E+7	0	US	itscaitlinhello
15/11/2022 23:16	162607	2260	4432	627859	She has a secret. And it's	movies, foryou, pov, f	well, well, well. if it is	7.16639E+7	7.16639E+7	0	US	itscaitlinhello
30/10/2022 18:59	79734	491	333	369475	Naked, WET. #fyp #foryou	halloween, cousins, fo	hey, sure, sure, you m	7.16038E+7	7.16038E+7	0	US	itscaitlinhello
13/10/2022 22:55	165465	1225	6129	791826	Good friend dinner with	kimkardashian, foryou	Jeff Bezos is like, wher	7.15413E+7	7.15413E+7	0	US	itscaitlinhello
11/10/2022 22:21	741690	11690	24469	2254431	Avery: “I know it's mornin	#fyp #foryou, parents, fyp	free can you please lo	7.15E+18	7.15338E+7	0	US	itscaitlinhello
08/10/2022 15:07	62819	636	1342	383807	Nicola Peltz-Beckham w	fyp, foryou, foryoupag	hi British Vogue it is m	7.15216E+7	7.15216E+7	0	US	itscaitlinhello

Note: The screenshot is intended solely to illustrate the workflow and data structure, not to expose raw data. Only a partial view is shown to minimise any risk of unnecessary disclosure.

Figure 6 Cleaned dataset from Manual Calibration built on Purchased data

Cleaned dataset from Manual Calibration built on Purchased data												
Date	Likes	Comments	Shares	Views	video_descriptions	hashtag	transcriptions	topics	Notes	covid	Url	MusicUrl
2022-12-28	1,300,000	11,400	13,000	7,700,000	I LOVE YOUR SON	fyp, foryou, foryoupag	hi. I know everyone's	relationship	glass, wine		https://www.https://v77Yes	148
2022-11-24	67,100	921	1,288	748,000	Happy thanksgiving, resp	fyp, foryou, foryoupag	hey, I know it's mornin	relationship	filter; sound effects		https://www.https://v77Yes	51
2022-11-15	160,700	1,745	1,736	1,100,000	She has a secret. And it's	fyp, foryou, foryoupag	well, well, well. if it is	film			https://www.https://v77Yes	97
2022-10-30	78,900	434	313	789,700	Naked, WET. #fyp #foryou	fyp, foryou, foryoupag	hey, sure, sure, you m	family			https://www.https://v77Yes	80
2022-10-13	164,400	1,044	2,065	1,500,000	Good friend dinner with	fyp, foryou, foryoupag	Jeff Bezos is like, wher	celebrity, TV show, friendship			https://www.https://v77Yes	70
2022-10-11	737,400	8,917	13,400	4,800,000	Avery: “I know it's mornin	#fyp #foryou #fyp, foryou, foryoupag	free can you please lo	social occasion (gathering), family			https://www.https://v77Yes	88
2022-10-08	62,000	530	544	665,200	Nicola Peltz-Beckham w	fyp, foryou, foryoupag	hi British Vogue it is m	celebrity, sunglasses, hair clip, gyn			https://www.https://v77Yes	143

Cleaned dataset from Manual Calibration built on Purchased data												
Date	Likes	Comments	Shares	Views	video_descriptions	hashtag	transcriptions	topics	Notes	covid	Url	MusicUrl
2022-12-28	1,300,000	11,400	13,000	7,700,000	I LOVE YOUR SON	fyp, foryou, foryoupag	hi. I know everyone's	relationship	glass, wine		https://www.https://v77Yes	148
2022-11-24	67,100	921	1,288	748,000	Happy thanksgiving, resp	fyp, foryou, foryoupag	hey, I know it's mornin	relationship	filter; sound effects		https://www.https://v77Yes	51
2022-11-15	160,700	1,745	1,736	1,100,000	She has a secret. And it's	fyp, foryou, foryoupag	well, well, well. if it is	film			https://www.https://v77Yes	97
2022-10-30	78,900	434	313	789,700	Naked, WET. #fyp #foryou	fyp, foryou, foryoupag	hey, sure, sure, you m	family			https://www.https://v77Yes	80
2022-10-13	164,400	1,044	2,065	1,500,000	Good friend dinner with	fyp, foryou, foryoupag	Jeff Bezos is like, wher	celebrity, TV show, friendship			https://www.https://v77Yes	70
2022-10-11	737,400	8,917	13,400	4,800,000	Avery: “I know it's mornin	#fyp #foryou #fyp, foryou, foryoupag	free can you please lo	social occasion (gathering), family			https://www.https://v77Yes	88
2022-10-08	62,000	530	544	665,200	Nicola Peltz-Beckham w	fyp, foryou, foryoupag	hi British Vogue it is m	celebrity, sunglasses, hair clip, gyn			https://www.https://v77Yes	143

As is typical for behavioural trace data, raw API output required extensive preprocessing. Technical constraints, including rate limits, pagination handling, and emoji decoding (Figure 3) [Insert (Figure3) here], and incomplete exposure of video-level features, restricted both the volume and granularity of extractable data. Every cleaning step was systematically documented, with derived variables cross-checked against supplementary metadata sources (Lazer et al., 2021; Salganik, 2018). This approach ensures transparency, reproducibility, and interpretability in subsequent modelling.

Finally, manual calibration was undertaken between April and July 2023, with videos verified and collected directly from TikTok.com. This step was designed to check whether there were missing videos between the actually available metadata on the website and the purchased source or API source. Therefore, it is reasonable to compare the number of videos from three sources. This served as a benchmark against which to evaluate the completeness and accuracy of both API- and analytics-based retrieval. Research-ready version presented in Figure 4 [Insert (Figure4) here].

3.3 Data Analyses

To evaluate the TikTok Research API's efficacy for computational social science, the analysis proceeded in two phases, addressing both macro-level data coverage and micro-level metric precision (Figure 5) [Insert (Figure5) here]. This dual approach aligns with RQ1 by assessing the API's technical capacity for multimodal data collection, and with RQ2 by examining whether systematic discrepancies across data sources could introduce demographic or algorithmic biases that compromise the fairness of societal insights. In other words, if certain groups or content types are systematically under- or over-represented due to data source effects, subsequent interpretations of societal trends risk being skewed.

Figure 7 Workflow diagram illustrating the two-phase data analysis process



Phase 1: Number of Videos

The first phase compared the monthly total number of videos (dependent variable) retrieved from the three data sources (API, manual calibration, and Analisa.io) to assess the API's ability to provide a comprehensive dataset. The raw dataset included 14 creators over 36 months, including two optional candidates (@steven_he and @nicholas_flannery). The initial specification included 12 creators, yielding a balanced panel of 1,251 observations.

After data harmonisation, linear mixed-effects models were fitted using the lme4 package in R, with creator gender (male/female), nationality (UK/US), and year of publication (2020, 2021, 2022) as control variables. A random intercept was added to account for creator-level clustering.

Robustness analyses were conducted in two subsets: (1) excluding four creators lacking purchased source data, which left 829 observations across the remaining eight creators. (2) excluding both the four creators with missing purchased data and one high-output outlier, yielding 733 observations across seven creators.

Phase 2: Quantity and Quality of Variables

The second phase examined data completeness and accuracy. First, the number of available metadata variables per video was compared across sources to provide a holistic measure of coverage. Next, the accuracy of key engagement metrics, like count, view count, comment count and share count, was assessed, as these are widely used indicators of content impact in social media research. Their reliability is critical:

systematic under- or over-reporting could distort comparative analyses and, by extension, the fairness of societal interpretations.

For comparability, only videos present in both sources were retained; those unique to a single source were excluded, thus leaving paired videos from 10 participants in total. Engagement analyses were restricted to API and manually calibrated data, as Analisa.io's rounded outputs closely mirrored manual counts, limiting their utility for precision testing. Notably, engagement metrics of @lizzza, @ameliadimz, @maddiegracejepson, and @hannahstocking from manual calibration are unavailable due to their lack of purchased source.

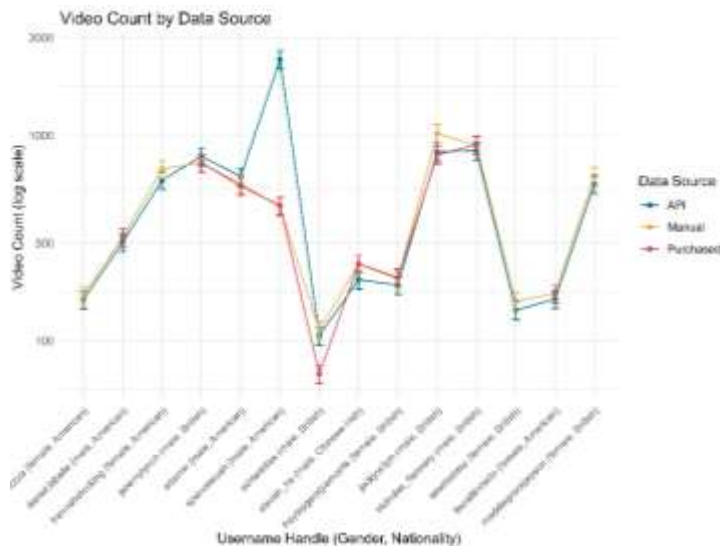
Linear mixed-effects models were also used, but with log-transformed outcome variables (via $\log_{1p}$) to normalise distributions. Fixed effects included data source, creator gender, nationality, and year of publication, while random intercepts for the creators and topics accounted for repeated measures. Topics were coded manually (e.g., childhood, family, friendship, see details in Appendix 5). Alongside quantitative checks, a qualitative review assessed the internal consistency of metadata fields.

Additionally, the newly introduced voice_to_text (transcription) variable was evaluated for validity, given its potential to introduce further demographic or linguistic bias. A binomial logistic regression was used to compare the differences in the proportion of videos with a transcription (dependent variable: whether a video had a transcription) as a function of gender and nationality (independent variables).

Preliminary Findings

4.1 Number of Videos

Figure 8 Video count by data source



Descriptively, on average, each month, the API ($M = 19.70, SD = 27.30$) and Manual sources ($M = 12.41, SD = 16.31$) generally captured a higher number of videos in raw counts than the Purchased source ($M = 11.56, SD = 16.70$). Figure 6 [Insert (Figure6) here] presents the total number of videos between 2020 and 2022 across three data sources, API, Manual, and Purchased, for 14 users, with each user identified by username, gender, and nationality. Heterogeneous correlation analysis indicated negligible associations among nationality, data source, and gender (all $|r| < .05$).

The linear mixed-effects analyses revealed consistent differences in estimated upload counts between data sources. In the full model, including all 14 creators, both the purchased dataset ($\beta = -7.98, p < 0.001, 95\% CI [-10.464, -5.523]$) and the manual dataset ($\beta = -4.80, p < 0.001, 95\% CI [-7.000, -2.620]$) had significantly lower counts than the API baseline. Gender was not a significant factor, but male creators tended to upload content more frequently. ($\beta = 12.80, p = 0.068, 95\% CI [-1.12, 26.73]$), while nationality effects were not significant ($\beta = -15.764, p = 0.235, 95\% CI [-37.899, 6.371]$).

When the two optional candidates (@steven_he and @nicholas_flannery) were excluded, the pattern persisted, with larger estimated differences for purchased ($\beta = -10.17, p < 0.001, 95\% CI [-12.950, -7.408]$) and manual ($\beta = -5.88, p < 0.001, 95\% CI [-8.279, -3.496]$) sources than API. Further exclusions of the four creators without purchased data ($n = 8, N = 829$) showed

that API-sourced monthly uploads were still higher than Purchased ($\beta = 11.764, p < 0.001, 95\% CI [8.560, 14.979]$), and Manual ($\beta = 9.034, p < 0.001, 95\% CI [5.802, 12.300]$).

As shown in Figure 6 [Insert (Figure6) here], there is an outlier of API (@spencewhuah). Thus, in a more restricted but balanced subset (e.g., excluding 4 creators absent from purchased data, 2 optional candidates, and the outlier, finally 7 creators), the negative coefficients for purchased and manual sources were smaller and statistically insignificant (e.g., purchased: $\beta = -1.20, p = 0.214, 95\% CI [-3.093, 0.683]$; manual: $\beta = -1.97, p = 0.937, 95\% CI [-1.970, 1.806]$), indicating the results' sensitivity to sample composition.

4.2 Quantity and Quality of Variables

The API generally provides a greater number of variables and with higher quality than alternative sources. Specifically, the API offers approximately twenty-two variables, including eight unique fields unavailable in other sources: effect ID, hashtag description, hashtag ID, playlist ID, region code, transcription, video label, and verified STEM information. Additionally, four variables (creation time, like count, music ID, and view count) exhibited greater precision (Table 4).

Nonetheless, the API has limitations; for example, it fails to decode emojis in video descriptions but showed “garbled text” or “mojibake” (Kita et al., 2022) when data are exported into CSV or Excel formats, requiring manual correction and extra online tools (Faust, 2017/2024) during cleaning. Furthermore, some variables (e.g., favourite count, hashtag description, video duration) were introduced incrementally after the API's initial release.

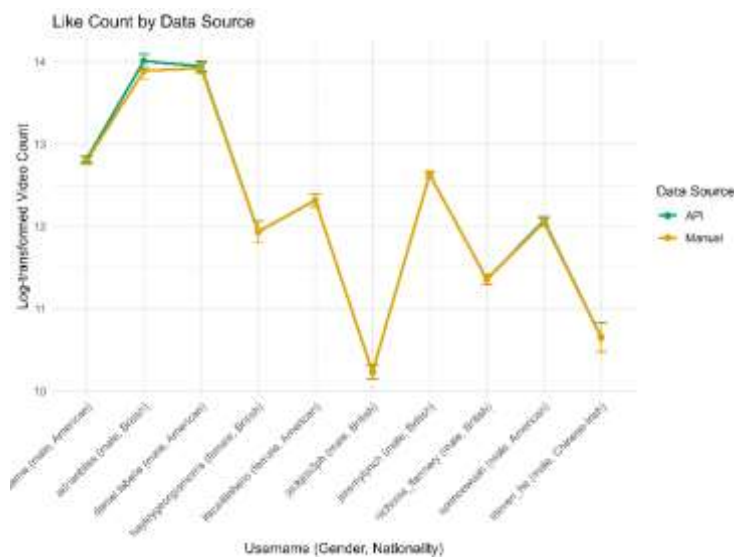
Table 4 Variable Quality from Three Sources

Content Made Available for Research								
Videos								
	Variable Names	Availability Research API	Analisa.io	Manual	Quality Research API	Analisa.io	Manual	Notes
1	comment count(comment_count)	Available	Available	Available	Accurate to the last digit	Accurate to the last digit	Accurate to the last digit	This is the total number of comments posted on a video.
2	create time (create_time)	Available	Available	Available	Accurate to the minute	Accurate to the minute	Accurate to date	This is the time when the video was created.
3	effect_ids	Available	Not Available	Not Available	Unique	NA	NA	The list of effects applied on the video.
4	favourite count (favourite_count)	Available after July 2024	Not Available	Available	Accurate to last digit	NA	Accurate to last digit	The number of favourites a video receives.
5	hashtags (hashtag_name)	Available	Available	Available	Accurate	Accurate	Accurate	The list of hashtags used in the video.
6	hashtag_description	Available after July 2024 or later	Not Available	Not Available	Unique	NA	NA	Returns a description for a hashtag_name if one exists.
7	hashtag_id	Available after July 2024 or later	Not Available	Not Available	Unique	NA	NA	Returns the unique hashtag_ids for each hashtag.
8	like count (like_count)	Available	Available	Available	Accurate to last digit	Rounded number to hundred	Rounded number to hundred	The total number of likes on a TikTok video, created by users by clicking the “Heart” icon.
9	music id (music_id)	Available	Available	Not Available	Accurate to last digit	Binary (original or else)	NA	This is the music_id used in the video.
10	playlist_id	Available	Not Available	Not Available	Unique	NA	NA	The ID of the playlist that the video belongs to.
11	region code	Available	Not Available	Not Available	Unique	NA	NA	A two-digit code for the country where the video creator registered their account.
12	share count (share_count)	Available	Available	Available	Accurate to last digit	Accurate to last digit	Accurate to last digit	The total number of times a TikTok video has been shared by clicking the “Share” button with the video.
13	transcription (voice_to_text)	Available	Not Available	Not Available	Unique (Low coverage)	NA	NA	Voice to text and subtitles (for videos that have voice to text features on, show the texts already generated)
14	username	Available	Available	Available	Accurate	Accurate	Accurate	This is the username of the video creator.
15	video description (video_description)	Available	Available	Available	Hard to decode emojis sometimes	Accurate	Accurate	This is the description of the video.
16	video duration in seconds (video_duration)	Available after July 2024	Available	Available	Accurate	Accurate	Accurate	By the end of my data collection (June 2024), Research API only offered the field name of “video_length”, which will only return “SHORT”, “MID”, “LONG”, “EXTRA_LONG”, so I did not include that index in the first place in the R codes. But I checked the official codebook in July, “video_duration” was added, and the precise length of the video can be returned through the query, along with another 2 field names of “is_stem_verified” and “favourites_count”
17	video id	Available	Not Available	Available	Accurate	NA	Accurate	The unique identifier of the TikTok video. This is also called “item_id” or “video_id”. This is a number that can be used to reconstruct the URL link to access the video.
18	video_label	Available after July 2024 or later	Not Available	Available	Unique	NA	NA	Returns any labels applied to a video such as “election labels” (Ex: Get info on the U.S elections)
19	video_mention_list	Available after July 2024 or later	Available	Available	Available	Available	Available	Returns the other tagged users in a video.
20	view count (view_count)	Available	Available	Available	Accurate to last digit	Rounded number to hundred	Rounded number to hundred	This is the total number of views for a video on TikTok.
21	whether verified stem information (is_stem_verified)	Available	Not Available	Available	Unique	NA	NA	Whether the video has been verified as being high quality STEM content.
22	video url	Partially Available	Available	Available	NA	Accurate	Accurate	Video URL can be constructed with video ID

Note: details of variable names can be found in Appendix 4

4.2.1 Variable Quality: Like Count

Figure 9 Monthly like count by data source (manual and API)



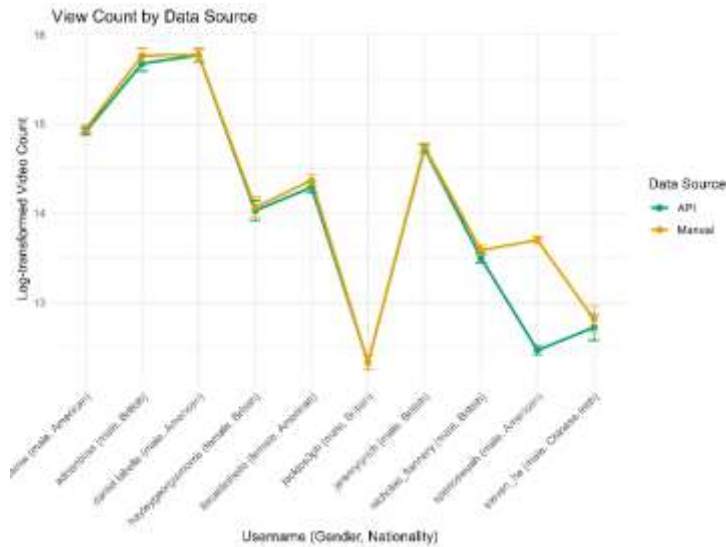
Across both manual and API sources ($N = 10$ creators, 8,424 observations), the average like count per observation was 5.47×10^5 ($SD = 1.23 \times 10^6$), with values ranging from 17 to 24.6 million. The mean like count was 5.50×10^5 ($SD = 1.23 \times 10^6$) for API entries and 5.45×10^5 ($SD = 1.22 \times 10^6$) for manual entries. Variance inflation factors indicated no evidence of problematic multicollinearity among predictors (all $GVIF^{(1/(2 \cdot Df))} \leq 1.03$). Figure 7 [Insert (Figure7) here] presents the monthly like counts with log-transformed values. In the model including all ten creators (excluding those without purchased data), the year of posting emerged as a significant predictor of like counts: relative to 2020, like counts were higher in 2021 ($\beta = 1.583$, $p < 0.001$, 95% CI [1.476, 1.690]) and 2022 ($\beta = 1.828$, $p < 0.001$, 95% CI [1.725, 1.931]), and 2022 was in turn even higher than 2021 ($\beta = 0.245$, $p < 0.001$, 95% CI [0.164, 0.326]). By contrast, there was no significant difference between manual and API sources ($\beta = 0.011$, $p = 0.737$, 95% CI [-0.051, 0.072]). Gender and nationality effects were not significant (p 's > 0.171).

When the two optional candidates were excluded, the pattern persisted. Data source effects remained negligible ($\beta = 0.013$, $p = 0.698$, 95% CI [-0.053, 0.079]), while year effects

were again substantial: 2021 ($\beta = 1.736$, $p < 0.001$, 95% CI [1.629, 1.844]) and 2022 ($\beta = 1.530$, $p < 0.001$, 95% CI [1.426, 1.633]) both exceeded 2020. There was no gender or nationality effects.

4.2.2 Variable Quality: View Count

Figure 10 Monthly view count by data source (manual and API)

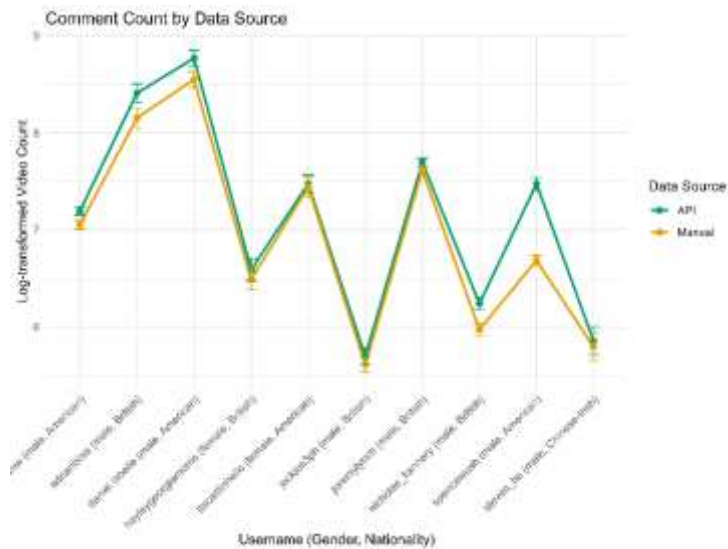


As for the view count ($N = 8,424$ observations, 10 creators), API-sourced data ($M = 3.69 \times 10^6$, $SD = 9.08 \times 10^6$) appeared to be lower than Manual ($M = 3.90 \times 10^6$, $SD = 9.17 \times 10^6$). As with the like-count model, predictors showed negligible intercorrelations (all $GVIF^{(1/(2 \cdot Df))} \leq 1.03$), confirming their suitability for inclusion. Figure 8 [Insert (Figure8) here] presents the monthly like counts with log-transformed values.

An analysis including all ten creators, showed that API-sourced counts were significantly fewer than manual counts ($\beta = -0.162$, $p < 0.001$, 95% CI [-0.218, -0.106]), while gender and nationality were not significant (p 's > 0.327). Additionally, view counts in 2022 were higher than in 2021 ($\beta = 0.181$, $p < 0.001$, 95% CI [0.108, 0.254]), the latter of which were in turn higher than in 2020 ($\beta = 1.573$, $p < 0.001$, 95% CI [1.477, 1.669]) When the two optional candidates were excluded, reducing the sample to eight creators, the pattern remained (see details in Appendix 6-A).

4.2.3 Variable Quality: Comment Count

Figure 11 Monthly comment count by data source (manual and API)

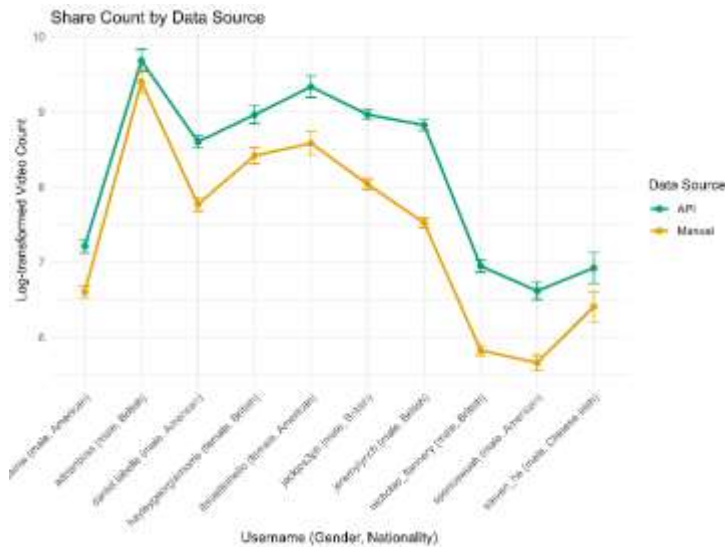


In terms of comment count ($N = 8,424$ observations, 10 creators), API-sourced data ($M = 4239.60, SD = 1.06 \times 10^4$) appeared to be higher than Manual ($M = 3375.56, SD = 9923.08$). Predictor variables were effectively independent (all $GVIF^{(1/(2 \cdot Df))} \leq 1.03$; categorical associations were weak, Cramér's $V < .12$). Figure 9 [Insert (Figure9) here] presents the monthly like counts with log-transformed values.

The data source predicted comment counts: API-sourced counts were significantly higher than manual counts ($\beta = 0.211, p < 0.001, 95\% CI [0.148, 0.275]$). While gender and nationality effects were not significant (p 's > 0.172), years were significant. Comment count in 2021 exhibited substantially higher values than 2020 ($\beta = 1.501, p < 0.001, 95\% CI [1.391, 1.610]$), but lower than 2022 ($\beta = -0.131, p = 0.002, 95\% CI [-0.215, -0.048]$). When the two optional candidates were excluded, reducing the sample to eight creators, the pattern remained (see details in Appendix 6-B).

4.2.4 Variable Quality: Share Count

Figure 12 Monthly share count by data source (manual and API)

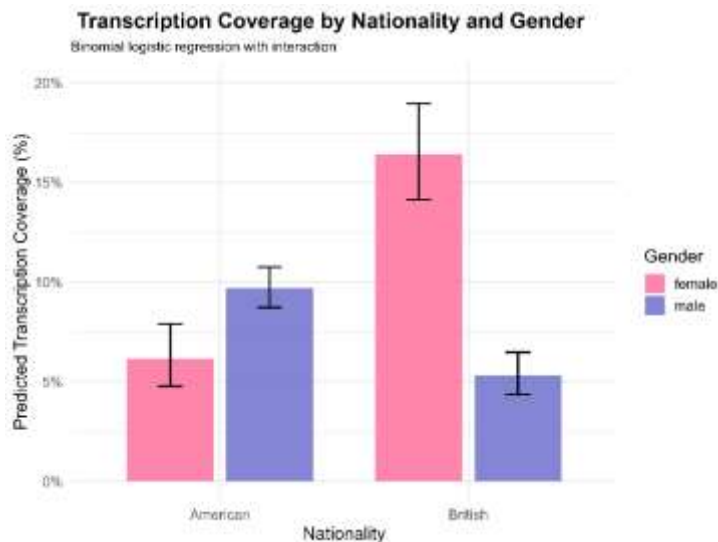


As for share count ($N = 8,424, M = 1.75 \times 10^4, SD = 9.09 \times 10^4$), videos collected via the API ($M = 2.85 \times 10^4, SD = 8.92 \times 10^4$) had a higher mean values than manual collection ($M = 1.35 \times 10^4, SD = 1.16 \times 10^5$). Consistent with previous models, no problematic multicollinearity was detected (all $GVIF^{1/(2 \cdot Df)} \leq 1.03$). Figure 10 [Insert (Figure10) here] details the monthly share counts with log-transformed values ($n=10$ creators).

Similar to comment counts, the analysis confirmed that API-sourced share counts were substantially higher than manual counts ($\beta = 0.922, p < 0.001, 95\% CI [0.840, 1.003]$), while gender and nationality were not significant ($p's > 0.229$). Additionally, relative to 2020, share counts were higher in 2021 ($\beta = 1.561, p < 0.001, 95\% CI [1.421, 1.701]$) and 2022 ($\beta = 0.529, p < 0.001, 95\% CI [0.394, 0.665]$), but 2022 were lower than 2021 ($\beta = -0.551, p < 0.001, 95\% CI [-0.678, -0.424]$).

4.2.5 Variable Quality: Transcription

Figure 13 Transcription rate of Research API



The API returns voice-to-text transcriptions, one of its unique fields, though only for approximately 10% of videos on average (Figure 11) [Insert (Figure11) here] out of API-driven dataset. A binomial logistic regression model ($n = 12$) was fitted to examine whether transcription percentages (i.e., the number of transcriptions divided by the number of videos) varied by creator nationality and gender, excluding two optional candidates (@steven_he and @nicholas_flannery). Main effects suggested that British creators had higher odds of transcription than American creators ($OR = 2.989, p < 0.001, 95\% CI [0.780, 14.22]$). Also, male creators had higher odds than female creators ($OR = 1.635, p < 0.001, 95\% CI [0.208, 12.793]$). However, there was a significant interaction between nationality and gender ($OR = 0.175, z = -8.54, p < .0001, 95\% CI = [-2.15, -1.35]$), indicating that the nationality effect differed markedly by gender.

Post hoc Tukey-adjusted pairwise comparisons showed that British females had significantly higher odds of providing transcriptions than American females (odds ratio. = 2.99, $p < 0.001, 95\% CI [0.220, 40.509]$) British females also had higher odds than American males (odds ratio = 1.83, $p < 0.001, 95\% CI [1.388, 2.409]$) and British males (odds ratio = 3.50, $p < 0.001, 95\% CI [2.441, 5.004]$). American males were more likely than American females to provide transcriptions (odds ratio. = 1.63, $p =$

0.0053, 95% *CI* [0.417, 0.897]), whereas British males did not differ significantly from American females ($p = 0.805$, 95% *CI* [0.748, 1.827]) British males, however, had significantly lower odds than both British females and American males ($p < 0.001$ In each case).

Discussion

This study has critically evaluated the TikTok Research API's capacity to support computational social science, addressing its effectiveness for multimodal data collection (RQ1), the manifestation of biases in retrieved data (RQ2), and the potential of multi-method integration to mitigate these challenges (RQ3). Drawing on a dataset of 6,373 videos collected via multiple methodologies, the findings depict a nuanced landscape: the API offers clear strengths, notably scalable, ethically compliant retrieval, yet these are offset by inconsistent metadata and systematic demographic biases.

The analysis reveals marked variability in video counts, engagement metrics, and transcription coverage across creators, with patterns strongly influenced by data source, nationality and gender for transcriptions. First, the Research API retrieves marginally higher video volumes than alternative methods whilst providing access to unique variables unavailable through manual collection or third-party services. However, this advantage proves uneven, as differences between data sources appear concentrated amongst specific creators rather than reflecting universal measurement biases. Second, while the data source and year exert a substantial influence on engagement metrics, gender and nationality do not. The consistent API advantage in engagement metrics raises questions about whether this reflects genuine behavioural differences or artefacts of the retrieval process.

Third, observed differences in transcription rates require careful interpretation. It is worth noting that TikTok's Automatic Captions feature was introduced in April 2021 (TikTok

Newsroom, 2021), midway through our study period (2020–2022). Because creators had the option to enable captions, this may explain why transcripts generated via the API appear predominantly after that launch date and are absent beforehand. The decision to activate voice-to-text functionality rests with individual creators rather than the API itself. However, the uneven distribution of transcripts across demographic groups suggests that uptake patterns may reflect broader systemic factors. It is plausible that American creators adopted this functionality at higher rates than their British counterparts, which could account for regional variation. Similarly, differential adoption rates between male and female creators may point to underlying inequalities in platform literacy, technical confidence, or community norms around accessibility features. Whilst the API faithfully returns whatever transcriptions creators have enabled, these creator-level choices may themselves be shaped by demographic and cultural contexts. This raises important questions about equity in platform feature adoption and the extent to which automated systems can perpetuate existing social disparities even when those systems function as designed.

These results carry important implications for future research employing multiple data-collection strategies in computational social science. While the API’s quantitative advantages are evident, they are compromised by technical inconsistencies and demographic skews that raise fundamental questions about validity and representativeness in societal trend analyses. Addressing these limitations will require methodological designs that integrate complementary sources, apply rigorous bias detection, and ensure transparency in reporting.

5.1 Multimodal Data Collection: Capabilities and Constraints (RQ1)

Quantitative Superiority and Scalability Advantages

The analysis indicates that the TikTok Research API is highly effective for large-scale data acquisition, outperforming purchased datasets and manual collection in quantitative terms.

Linear mixed-effects models confirm that it retrieves more videos overall, though results

remain sensitive to sample composition. Meanwhile, the study suggests that API-sourced data systematically capture higher viewing, commenting and sharing activity. Crucially, its advantage extends beyond volume: the API grants access to unique variables — including transcriptions, effect identifiers, and precise metadata fields—that remain unavailable through alternative approaches. These affordances enable rapid assembly of large longitudinal datasets, facilitating both quantitative modelling and qualitative content analysis.

This scalability marks a significant advance for computational social science. It enables the rapid assembly of datasets that would be prohibitively resource-intensive to collect manually. The ability to retrieve historical content spanning multiple years further enhances its value for longitudinal studies, particularly those examining temporal dynamics during crisis periods such as the COVID-19 pandemic. These capabilities support a wide range of analytical approaches, from quantitative modelling of engagement metrics to qualitative content analysis.

Technical Constraints Undermining Data Validity

However, these strengths are tempered by technical shortcomings that compromise validity. Unstable API performance, consistent with Ruz et al. (2023), produces inconsistent data volumes across collection periods. Content visibility and access issues arise through deletion or privacy settings, rendering content inaccessible via the API (TikTok for Developers, n.d.), whilst manual calibration remains limited to publicly available videos. The outlier “@spencewuah” illustrates a marked discrepancy in video counts between API and manual sources.

Further distortions occur because the API includes archived videos, such as drafts, inflating counts relative to platform-visible data. More data returned by the API does not necessarily represent better quality but may compromise results when researchers focus on platform-accessible content. Transcription coverage is particularly sparse; voice-to-text data exist for

around 10% of videos, predominantly from 2021 onwards. This temporal skew excludes much of the early pandemic period from linguistic analysis, potentially obscuring shifts in discourse during initial lockdowns.

Metadata inconsistencies, notably the API's failure to decode emojis, produce mojibake (Kita et al., 2022) and require manual correction (Faust, 2017/2024). As Cohn's foundational work on the visual language of sequential images (2013) demonstrates, pictorial elements, whether in comics or digital media, operate within structured semiotic systems. Subsequent research on emoji sequencing (Cohn et al., 2019) extends this framework, showing that emoji, like other visual modalities, often lack fully developed grammatical structures yet still interact systematically with accompanying text. Such issues challenge Tromble's (2021) assertion that minimal programming suffices for API research, as specialist intervention is often essential. Additionally, the Research API cannot return or store video or audio data, which must be obtained through third-party services (see Appendix 7 for a list of possible services). Prosody, captured through speech rhythm, intonation, and emphasis, offers cues to stance and affect otherwise absent from purely textual analysis (Arvaniti, 2020; Mauchand et al., 2020). In multimodal contexts, prosodic contours synchronise with gesture, jointly signalling pragmatic meanings such as information status, stance, and (im)politeness (Brown & Prieto, 2021; Kendon, 2004; McNeill, 2005). This integration supports the view that prosody and gesture operate as sister systems in sociopragmatic marking (Ambrazaitis & House, 2017; Prieto & Espinal, 2020). In short-form video, these systems work in concert with visual symbols, including emojis, to form a tightly integrated semiotic ensemble.

Ethical Constraints and Access Limitations

Regional restrictions shape content availability through government bans, platform moderation, and algorithmic filtering (Zhao, 2024). In this study, some U.S.-origin videos were inaccessible from the U.K., reducing dataset completeness and constraining

cross-national comparison. These constraints expose the API's limitations for longitudinal analysis and point to the need for multi-method calibration to safeguard validity. The API can increase data volume, but doing so demands substantial computational resources for processing and verification.

Technical and ethical constraints compound these issues. Strict retrieval limits and the requirement for formal ethical approval impede reproducibility, echoing earlier critiques of API-based research (Freelon, 2018). Access policies that privilege university-affiliated researchers raise equity concerns and appear misaligned with the EU's Digital Services Act commitment to transparency (Burnnat & Davidson, 2025). Challenges in data dissemination and retention derived from ToS still haunt open science (Bak-Coleman, 2023; Venkatagiri, 2023). Collectively, these barriers not only narrow the empirical scope but also shape the wider conditions under which regulated digital research is conducted.

5.2 Algorithmic and Demographic Biases: Compromising Research Equity (RQ2)

The analysis identifies pronounced disparities within TikTok Research API that compromise research equity. Transcription coverage systematically varies across demographic groups, with American and male creators demonstrating higher rates than their U.K. and female counterparts. Whilst the choice to enable TikTok's Automatic Captions feature rests with individual creators, the observed patterns suggest that adoption may be influenced by broader structural factors.

Several mechanisms may account for these disparities. American creators may have adopted auto-captioning functionality at higher rates than British creators, potentially reflecting earlier exposure to the feature, stronger platform literacy, or differing community norms around accessibility. Similarly, gendered differences in transcription availability may stem from variations in technical confidence, awareness of accessibility tools, or expectations regarding content presentation. Research on digital inequalities has consistently demonstrated that

platform feature adoption is rarely neutral but often reflects pre-existing social stratifications (Feng et al., 2021; Markl, 2022; Sari et al., 2021).

Moreover, even when creators do enable captions, the accuracy of automated speech recognition systems may vary systematically across demographic groups. ASR models are typically trained on datasets that over-represent particular dialects, accents, or demographic groups, leading to differential performance (Wassink et al., 2022). Creators whose speech patterns align with the dominant training data enjoy greater transcription accuracy. At the same time, those outside these norms may experience lower-quality outputs, potentially discouraging further use of the feature. This creates a feedback loop wherein initial technical disparities become reinforced through user behaviour.

These demographic skews have significant implications for the fairness and validity of societal trend analyses. Unrepresentative samples risk distorting interpretations of cultural phenomena, particularly in policy-relevant contexts where findings might inform interventions. In such cases, conclusions may disproportionately reflect the perspectives of privileged groups, limiting generalisability in global or gender-diverse settings. The marked attenuation of effects following outlier removal further indicates that extreme observations can disproportionately influence perceived differences between data collection methods. This underscores the importance of examining distributional properties and identifying influential cases before drawing inferences about data quality.

5.3 Multi-Method Integration: Mitigating Inconsistencies for Enhanced Validity (RQ3)

In light of the challenges identified in RQ1 and RQ2, the analysis demonstrates that integrating multiple data collection methods, combining API-derived data with manual calibration and third-party analytics, provides a robust means of improving validity. This triangulated approach mitigates inconsistencies, such as video-count discrepancies, and enhances completeness through cross-validation of engagement metrics. In doing so, it

addresses both technical limitations and bias-related distortions, thereby strengthening the evidential base for subsequent interpretation.

The comparative analysis reveals systematic differences in how engagement metrics are captured across methods. The data source has a significant influence on reported values: videos obtained via the API exhibit lower log-transformed view counts than those gathered manually, yet higher comment and share counts. This asymmetry suggests that platform-specific definitions or calculation procedures for engagement metrics may differ between collection routes. In other words, the same video can appear to perform differently depending on the method used, a non-trivial complication for comparative research.

Multi-method integration also proves critical for examining temporal engagement patterns. The year of posting emerged as a consistent predictor across metrics, with videos from 2021 recording markedly higher engagement than 2020 content. It is reasonable that established creators benefit from a larger reach over time, also known as “cumulative advantage” (De Oliveira Santini et al., 2020). These temporal shifts underscore the importance of accounting for contextual factors when interpreting engagement trends. Encouragingly, as of 15 August 2025, TikTok Research API includes a batch compliance task¹⁰ feature (TikTok for Developers, 2025), allowing researchers to submit lists of video or comment identifiers for validation. This process enables verification of whether each identifier remains publicly accessible, supporting accurate reporting of dataset availability at the time of analysis.

By incorporating alternative transcription services, such as Happy Scribe (see Appendix 7 for a list of possible services), and cross-checking outputs, this methodological pluralism reduces the impact of incomplete coverage and demographic skew. Crucially, it enables researchers to distinguish findings that are robust across methods from those that are artefacts of a particular

¹⁰ How to create a batch compliance task: https://developers.tiktok.com/doc/batch-compliance-apis?enter_method=left_navigation

collection strategy. In periods of rapid social change, when data quality assumes heightened importance, such integration is not merely desirable but essential for producing credible, equitable, and generalisable insights.

5.4 Data Accessibility, Methodological Transparency, and Ethical Considerations

The empirical patterns identified in RQ1–RQ3 reinforce the existing literature on the tension between data legitimacy and research reproducibility in computational social science (Davidson et al., 2023), including demographic and algorithmic biases, metric inconsistencies across different collection methods, and the benefits of multi-method integration. While the API enables structured, large-scale data retrieval, its utility is constrained by platform-imposed rate limits and post-hoc deletions by content creators. Such constraints, consistent with Ruz et al. (2023) and Entrena-Serrano et al. (2025), undermine data completeness, particularly in longitudinal studies of temporally sensitive phenomena such as pandemic-related humour. Echoing Pfeffer et al. (2018), sample-streaming endpoints also exhibit non-random omissions, manipulation, and popularity bias, necessitating corrective measures such as stratified sampling or weighting to preserve representativeness.

Discrepancies in engagement metrics further complicate interpretation. API-sourced videos tend to report lower view counts but higher comment and share counts than their manually collected equivalents, aligning with Pearson et al. (2025), who show that both data source and year significantly influence TikTok viewership. Notably, no significant effects emerged for data source in like-count analyses, suggesting that likes may be more consistently measured across methods. These asymmetries highlight the need for methodological transparency when comparing engagement metrics derived from different pipelines.

Ethical considerations compound these methodological challenges. The API's exclusivity disadvantages independent scholars and civil society groups, reinforcing structural inequities in research access. Moreover, its granular engagement data, including comment histories,

raise privacy risks such as potential deanonymisation, even within TikTok's consent frameworks (Mimizuka et al., 2025). These risks are magnified in high-stakes contexts, such as public health or political discourse, where the digital divide exacerbates disparities in who can produce, access, and analyse data (Bezjak et al., 2018; Rossi & Lenzini, 2020). Robust ethical protocols and transparent governance are therefore imperative to balance analytical utility with accountability.

By leveraging a large, balanced dataset spanning two nations and genders, this study addresses the limitations of prior TikTok research, which often relied on smaller, non-systematic samples (e.g., Cervi & Divon, 2023). It complements emerging machine-learning applications of the TikTok Research API in political and behavioural domains (Corso et al., 2024; Pearson et al., 2025), while applying a critical social-science lens to issues of methodological rigour and representational equity. The findings affirm the API's potential as a scalable, compliant tool for computational social science, but also underscore the necessity of multi-method strategies to mitigate its shortcomings. Balancing its strengths (volume, unique variables) against its weaknesses (quality, bias) provides a framework for best practice in ensuring validity and fairness, thereby contributing to the ongoing discourse on platform-based methodologies.

5.5 Study Limitations

Whilst this study offers valuable insights into TikTok Research API for social science research, several limitations must be acknowledged. First, the restricted sample of creators poses a notable constraint. With data drawn from only 14 creators based in the United States and the United Kingdom, the findings may not fully represent the diverse TikTok creator population. This limitation restricts the generalisability of the results, particularly for creators from other regions or with varying audience demographics (Herring, 2010; Lomborg & Bechmann, 2014).

Second, the discrepancies between API and manual coding are unlikely to be attributable solely to time accumulation. Although the API data were collected several months after the manual coding, potentially inflating shares and comments by allowing more time to accumulate, uploads within a fixed past period are not changed; yet, the API still reports more uploads (suggesting manual under-coverage rather than a timing artefact). Moreover, despite this timing advantage, the API shows fewer views, indicating that view differences are not explained by timing alone and likely reflect differences in coverage or measurement. Additionally, the TikTok Research API introduced challenges affecting data quality. Transcriptions were available for approximately 10% of videos, limiting the scope of linguistic analysis. Inconsistent metadata accuracy, such as discrepancies in engagement metrics, further complicates the reliability of the findings. Finally, methodological constraints warrant consideration. The study's timeframe (2020–2022) may not reflect longer-term trends, whilst the integration of manual calibration with API data risks introducing inconsistencies. These limitations suggest caution in interpreting the results and highlight avenues for future research, such as expanding the sample diversity and refining data collection methods.

Conclusion

This study set out to interrogate the reliability, representativeness, and ethical dimensions of TikTok Research API, using a large, balanced dataset to examine how methodological choices shape the validity of social media research. By systematically addressing RQ1–RQ3, it has shown that while the API offers unprecedented opportunities for structured, large-scale analysis, its outputs are neither neutral nor complete. Algorithmic and demographic biases, metric inconsistencies, and coverage gaps are not peripheral flaws but structural features of the data environment, features that, if left unexamined, risk distorting scholarly narratives and policy-relevant insights. Yet, these limitations are tractable: multi-method integration, bias-

aware modelling, and transparent reporting can materially improve both the equity and reproducibility of platform-based research.

This critical assessment highlights the API's dual character. On the one hand, it provides compliant and scalable access to multimodal data, offering a legally sanctioned means of integrating quantitative and qualitative analyses (Nguyen & Diederich, 2023). On the other hand, it exhibits technical instability, systematic biases towards American and male creators, and restrictive ethical constraints. Whilst the API enhances computational social science by enabling diverse analytical approaches, its shortcomings in data quality, transparency, and accessibility necessitate a multi-method strategy to ensure valid and equitable insights. By doing so, this research contributes to ongoing debates on platform governance and methodological pluralism, demonstrating how comparative strategies can mitigate the deficiencies of single-source datasets. Enhancing the reliability and representativeness of social media data remains essential for robust societal understanding and for fostering a more equitable digital research ecosystem.

Future work should adopt integrated approaches, combining API data with manual scraping or third-party platforms to cross-validate findings and address coverage gaps. Documenting discrepancies enhances transparency (Bezjak et al., 2018), while post-weighting adjustments can counter demographic biases (Ohme et al., 2024). Platform developers, for their part, should increase retrieval limits, expand transcription coverage, clarify metric definitions, and broaden access beyond academic institutions to promote inclusivity.

Finally, these findings open pathways for comparative API studies, for example, across Douyin, YouTube Shorts, and Instagram Reels, and for deeper examination of transcription biases and algorithmic opacity, particularly in relation to misinformation and political mobilisation. Leveraging automated tools or language models offers further potential,

provided ethical standards are upheld. In this way, social media research can evolve to meet contemporary challenges while remaining methodologically rigorous, transparent, and fair.

References (Chapter 1):

- Acker, A., & Kreisberg, A. (2020). Social media data archives in an API-driven world. *Archival Science*, 20(2), 105–123. <https://doi.org/10.1007/s10502-019-09325-9>
- Agrawal, E. (2024). Going viral: An analysis of advertising of technology products on TikTok. arXiv. <https://doi.org/10.48550/ARXIV.2402.00010>
- Alexandre, I., Jai-sung Yoo, J., & Murthy, D. (2022). Make Tweets Great Again: Who Are Opinion Leaders, and What Did They Tweet About Donald Trump? *Social Science Computer Review*, 40(6), 1456–1477. <https://doi.org/10.1177/08944393211008859>
- Ambrazaitis, G., & House, D. (2017). Multimodal prominences: Exploring the patterning and usage of focal pitch accents, head beats and eyebrow beats in Swedish television news readings. *Speech Communication*, 95, 100–113. <https://doi.org/10.1016/j.specom.2017.08.008>
- Arora, A., Bansal, S., Kandpal, C., Aswani, R., & Dwivedi, Y. (2019). Measuring social media influencer index- insights from Facebook, Twitter and Instagram. *Journal of Retailing and Consumer Services*, 49, 86–101. <https://doi.org/10.1016/j.jretconser.2019.03.012>
- Arvaniti, A. (2020). The phonetics of prosody. In A. Arvaniti, *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780199384655.013.411>
- Bak-Coleman, J. (2023, February 22). TikTok’s API guidelines are a minefield for researchers | TechPolicy.Press. Tech Policy Press. <https://techpolicy.press/tiktoks-api-guidelines-are-a-minefield-for-researchers>
- Bezjak, S., Clyburne-Sherin, A., Conzett, P., Fernandes, P., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., Psomopoulos, F., Ross-Hellauer, T., Schneider, R., Tennant, J., Verbakel, E., Brinken, H., & Heller, L. (2018). Open science training handbook. Zenodo. <https://doi.org/10.5281/zenodo.1212496>

- Brown, L., & Prieto, P. (2021). Gesture and prosody in multimodal communication. In M. Haugh, D. Z. Kádár, & M. Terkourafi (Eds.), *The Cambridge Handbook of Sociopragmatics* (1st ed., pp. 430–453). Cambridge University Press.
<https://doi.org/10.1017/9781108954105.023>
- Bruns, A. (2019). After the ‘APIcalypse’: Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11), 1544–1566. <https://doi.org/10.1080/1369118X.2019.1637447>
- Burnat, F. A. D., & Davidson, B. I. (2025). The Accountability Paradox: How Platform API Restrictions Undermine AI Transparency Mandates. arXiv.
<https://doi.org/10.48550/ARXIV.2505.11577>
- Cervi, L., & Divon, T. (2023). Playful activism: Memetic performances of Palestinian resistance in TikTok #challenges. *Social Media + Society*, 9(1), 20563051231157607.
<https://doi.org/10.1177/20563051231157607>
- Chang, A. (2018, March 23). The Facebook and Cambridge Analytica scandal, explained with a simple diagram. Vox. <https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram>
- Chen, Z., & Zhang, Q. (2021). A survey study on successful marketing factors for Douyin (TikTok). In F. F.-H. Nah & K. Siau (Eds.), *HCI in Business, Government and Organizations* (pp. 22–42). Springer International Publishing.
https://doi.org/10.1007/978-3-030-77750-0_2
- Cheng, Z., & Li, Y. (2024). Like, comment, and share on TikTok: Exploring the effect of sentiment and second-person view on the user engagement with Tiktok news videos. *Social Science Computer Review*, 42(1), 201–223.
<https://doi.org/10.1177/08944393231178603>

- Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. Bloomsbury Academic, An imprint of Bloomsbury Pub. Plc.
- Cohn, N., Engelen, J., & Schilperoord, J. (2019). The grammar of emoji? Constraints on communicative pictorial sequencing. *Cognitive Research: Principles and Implications*, 4(1), 33. <https://doi.org/10.1186/s41235-019-0177-0>
- Corso, F., Pierri, F., & De Francisci Morales, G. (2024). What we can learn from TikTok through its Research API. *Companion Proceedings of the 16th ACM Web Science Conference*, 110–114. <https://doi.org/10.1145/3630744.3663611>
- Davidson, B. I., Wischerath, D., Racek, D., Parry, D. A., Godwin, E., Hinds, J., van der Linden, D., Roscoe, J. F., Ayravainen, L., & Cork, A. G. (2023). Platform-controlled social media APIs threaten open science. *Nature Human Behaviour*, 7(12), 2054–2057. <https://doi.org/10.1038/s41562-023-01750-2>
- De Oliveira Santini, F., Ladeira, W. J., Pinto, D. C., Herter, M. M., Sampaio, C. H., & Babin, B. J. (2020). Customer engagement in social media: A framework and meta-analysis. *Journal of the Academy of Marketing Science*, 48(6), 1211–1228. <https://doi.org/10.1007/s11747-020-00731-5>
- Entrena-Serrano, C. (2025). Watch, scroll, repeat: How interface design shapes consumptive curation affordances on Tiktok. *Social Media + Society*, 11(3), 20563051251358529. <https://doi.org/10.1177/20563051251358529>
- Entrena-Serrano, C., Degeling, M., Romano, S., & Çetin, R. B. (2025). TikTok’s research API: Problems without explanations. *arXiv*. <https://doi.org/10.48550/ARXIV.2506.09746>
- European Centre for Algorithmic Transparency. (2025). FAQs: DSA data access for researchers—European Commission. European Commission; European Commission.

https://algorithmic-transparency.ec.europa.eu/news/faqs-dsa-data-access-researchers-2025-07-03_en.

- Faust, W. (2024). Wolfgang42/ftfyweb [Smarty]. <https://github.com/wolfgang42/ftfyweb> (Original work published 2017)
- Feng, S., Kudina, O., Halpern, B. M., & Scharenborg, O. (2021). Quantifying bias in automatic speech recognition. arXiv. <https://doi.org/10.48550/ARXIV.2103.15122>
- Freelon, D. (2018). Computational Research in the Post-API Age. *Political Communication*, 35(4), 665–668. <https://doi.org/10.1080/10584609.2018.1477506>
- Haime, Z., & Biddle, L. (2025). Exploring mental health content moderation and well-being tools on social media platforms: Walkthrough analysis. *JMIR Human Factors*, 12, e69817. <https://doi.org/10.2196/69817>
- Herring, S. C. (2009). Web content analysis: Expanding the paradigm. In J. Hunsinger, L. Klastrup, & M. Allen (Eds.), *International Handbook of Internet Research* (pp. 233–249). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9789-8_14
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.
- Kirkpatrick, C. E., & Lawrie, L. L. (2024). TikTok as a source of health information and misinformation for young women in the United States: Survey study. *JMIR Infodemiology*, 4, e54663. <https://doi.org/10.2196/54663>
- Kita, H., Kitamura, Y., Hioki, H., Sakai, H., & Lin, D. (2022). The practice of basic informatics 2022. 1–236. <https://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/289713>
- Klug, D., Qin, Y., Evans, M., & Kaufman, G. (2021). Trick and please. A mixed-method study on user assumptions about the tiktok algorithm. *13th ACM Web Science Conference 2021*, 84–92. <https://doi.org/10.1145/3447535.3462512>
- Lachief. (2023). Analisa.io. Lachief. <https://www.lachief.io/tool-posts/analisa-io>

- Lazer, D., Hargittai, E., Freelon, D., Gonzalez-Bailon, S., Munger, K., Ognyanova, K., & Radford, J. (2021). Meaningful measures of human society in the twenty-first century. *Nature*, 595(7866), 189–196. <https://doi.org/10.1038/s41586-021-03660-7>
- Leedham, M., Lillis, T., & Twiner, A. (2020). Exploring the core ‘preoccupation’ of social work writing: A corpus-assisted discourse study. *Journal of Corpora and Discourse Studies*, 3(0). <https://doi.org/10.18573/jcads.26>
- Liang, X. (2021). Research on how to perceive their behavior for international high school students based on using tiktok with semi-structured interview. 796–799. <https://doi.org/10.2991/assehr.k.210407.151>
- Literat, I. (2021). “Teachers Act Like We’re Robots”: TikTok as a Window Into Youth Experiences of Online Learning During COVID-19. *AERA Open*, 7, 233285842199553. <https://doi.org/10.1177/2332858421995537>
- Lomborg, S., & Bechmann, A. (2014). Using APIs for Data Collection on Social Media. *The Information Society*, 30(4), 256–265. <https://doi.org/10.1080/01972243.2014.915276>
- Lu, Y., & (Cindy) Shen, C. (2023). Unpacking multimodal fact-checking: Features and engagement of fact-checking videos on Chinese TikTok (Douyin). *Social Media + Society*, 9(1). <https://doi.org/10.1177/20563051221150406>
- Markl, N. (2022). Language variation and algorithmic bias: Understanding algorithmic bias in British English automatic speech recognition. 2022 ACM Conference on Fairness Accountability and Transparency, 521–534. <https://doi.org/10.1145/3531146.3533117>
- Mauchand, M., Vergis, N., & Pell, M. D. (2020). Irony, prosody, and social impressions of affective stance. *Discourse Processes*, 57(2), 141–157. <https://doi.org/10.1080/0163853X.2019.1581588>
- McNeill, D. (2005). *Gesture and thought*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226514642.001.0001>

- Mimizuka, K., Brown, M. A., Yang, K.-C., & Lukito, J. (2025). Post-Post-API Age: Studying Digital Platforms in Scant Data Access Times. arXiv.
<https://doi.org/10.48550/ARXIV.2505.09877>
- Moir, A. (2023). The use of TikTok for political campaigning in Canada: The case of Jagmeet Singh. *Social Media + Society*, 9(1), 20563051231157604.
<https://doi.org/10.1177/20563051231157604>
- Mosnar, M., Skurla, A., Pecher, B., Tibensky, M., Jakubcik, J., Bindas, A., Sakalik, P., & Srba, I. (2025). Revisiting algorithmic audits of tiktok: Poor reproducibility and short-term validity of findings. *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3357–3366.
<https://doi.org/10.1145/3726302.3730293>
- Nguyen, H., & Diederich, M. (2023). How civil are comments on TikTok’s educational videos? Insights for learning at scale. *Proceedings of the Tenth ACM Conference on Learning @ Scale*, 292–296. <https://doi.org/10.1145/3573051.3596174>
- Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology*, 21(19), 1641–1646. <https://doi.org/10.1016/j.cub.2011.08.031>
- Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2024). Digital trace data collection for social media effects research: APIs, data donation, and (Screen) tracking. *Communication Methods and Measures*, 18(2), 124–141. <https://doi.org/10.1080/19312458.2023.2181319>
- Parikh, N. (2025, December 19). How to increase tiktok engagement: 16 expert insights. Socialinsider Blog; Socialinsider. <https://www.socialinsider.io/blog/how-to-get-more-engagement-on-tiktok/>

- Parikh, N. (2025, December 19). How to increase TikTok engagement: 16 expert insights. Socialinsider Blog; Socialinsider. <https://www.socialinsider.io/blog/how-to-get-more-engagement-on-tiktok/>
- Parisi, L., Mulargia, S., Comunello, F., Bernardini, V., Bussoletti, A., Nisi, C. R., Russo, L., Campagna, I., Lanfranchi, B., Croci, I., Grassucci, E., & Gesualdo, F. (2023). Exploring the vaccine conversation on TikTok in Italy: Beyond classic vaccine stances. *BMC Public Health*, 23(1), 880. <https://doi.org/10.1186/s12889-023-15748-y>
- Pascual, C. (2020). R API Tutorial: Getting Started with APIs in R. Dataquest. <https://www.dataquest.io/blog/r-api-tutorial/>.
- Pearson, G.D.H., Silver, N.A., Robinson, J.Y., Azadi, M., Schillo, B.A. and Kreslake, J.M. 2025. Beyond the margin of error: a systematic and replicable audit of the TikTok research API. *Information, Communication & Society*. 28(3), pp.452–470.
- Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, 20(10), 624–634. <https://doi.org/10.1038/s41583-019-0202-9>
- Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental imagery: Functional mechanisms and clinical applications. *Trends in Cognitive Sciences*, 19(10), 590–602. <https://doi.org/10.1016/j.tics.2015.08.003>
- Perriam, J., Birkbak, A., & Freeman, A. (2020). Digital methods in a post-API environment. *International Journal of Social Research Methodology*, 23(3), 277–290. <https://doi.org/10.1080/13645579.2019.1682840>
- Pfeffer, J., Mayer, K. and Morstatter, F. (2018). Tampering with Twitter’s Sample API. *EPJ Data Science*. 7(1), p.50.

- Pinto, G., Bickham, C., Salkar, T., Luceri, L., & Ferrara, E. (2024a). Tracking the 2024 us presidential election chatter on tiktok: A public multimodal dataset. arXiv. <https://doi.org/10.48550/ARXIV.2407.01471>
- Pinto, G., Burghardt, K., Lerman, K., & Ferrara, E. (2024b). Get-tok: A genai-enriched multimodal tiktok dataset documenting the 2022 attempted coup in Peru. arXiv. <https://doi.org/10.48550/ARXIV.2402.05882>
- Prieto, P., & Espinal, M. T. (2020). Negation, prosody, and gesture. In V. Déprez & M. T. Espinal (Eds.), *The Oxford Handbook of Negation* (1st ed., pp. 677–693). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198830528.013.34>
- Reisach, U. (2021). The responsibility of social media in times of societal and political manipulation. *European Journal of Operational Research*, 291(3), 906–917. <https://doi.org/10.1016/j.ejor.2020.09.020>
- Rejeb, A., Rejeb, K., Appolloni, A., Treiblmaier, H., & Iranmanesh, M. (2024). Mapping the scholarly landscape of TikTok (Douyin): A bibliometric exploration of research topics and trends. *Digital Business*, 4(1), 1–23. <https://doi.org/10.1016/j.digbus.2024.100075>
- Rogers, R., & Zhang, X. (2024). The Russia–Ukraine war in Chinese social media: LLM analysis yields a bias toward neutrality. *Social Media + Society*, 10(2), 20563051241254379. <https://doi.org/10.1177/20563051241254379>
- Rossi, A., & Lenzini, G. (2020). Transparency by design in data-informed research: A collection of information design patterns. *Computer Law & Security Review*, 37, 105402. <https://doi.org/10.1016/j.clsr.2020.105402>
- Ruz, S. S., Degeling, M., & Meßmer, K. (2023, October 16).  The Research API falls woefully short | auditing TikTok. Auditing TikTok. <https://tiktok-audit.com/blog/2023/the-TikTok-research-API-falls-woefully-short/>

- Salganik, M. J. (2018). *Bit by bit: Social research in the digital age*. Princeton University Press.
- Sari, L., Hasegawa-Johnson, M., & Yoo, C. (2021). Counterfactually fair automatic speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29, 3515–3525. <https://doi.org/10.1109/taslp.2021.3126949>
- Sato, M. (2023, February 21). Researchers will get access to TikTok data-Pending company approval. *The Verge*. <https://www.theverge.com/2023/2/21/23604737/tiktok-research-api-expansion-public-user-data-transparency>
- Sharma, K., Ferrara, E., & Liu, Y. (2022). Characterizing Online Engagement with Disinformation and Conspiracies in the 2020 U.S. Presidential Election. *Proceedings of the International AAAI Conference on Web and Social Media*, 16, 908–919. <https://doi.org/10.1609/icwsm.v16i1.19345>
- Stevens, G., O'Donnell, V. L., & Williams, L. (2015). Public domain or private data? Developing an ethical approach to social media research in an inter-disciplinary project. *Educational Research and Evaluation*, 21(2), 154–167. <https://doi.org/10.1080/13803611.2015.1024010>
- Sweney, M. (2022, October 4). TikTok reports \$1bn turnover across international markets. *The Guardian*. <https://www.theguardian.com/technology/2022/oct/04/tiktok-reports-1bn-turnover-across-international-markets>
- Thole, A. N. (2022). TikTok forensic scraper to retrieve user video details. 3998505 Bytes. <https://doi.org/10.25394/PGS.21676727.V1>
- TikTok (2024a). Digital Services Act: Publishing our second transparency report on content moderation in Europe. Newsroom | TikTok. <https://newsroom.tiktok.com/en-eu/digital-services-act-second-transparency-report>.

TikTok (2024b). Terms of Service | TikTok. TikTok.

<https://www.tiktok.com/legal/page/eea/terms-of-service/en>.

TikTok for Developers (2024). Research API | TikTok for Developers.

<https://developers.tiktok.com/products/research-api/>.

TikTok Newsroom. (2021, April 6). Introducing auto captions. Newsroom | TikTok;

ByteDance. <https://newsroom.tiktok.com/introducing-auto-captions?lang=en-GB>

Transparency Centre (2024). Publishing with Meta Content Library and Content Library API.

Meta for Developers. <https://developers.facebook.com/docs/content-library-and-api/citations/>.

Tromble, R. (2021). Where Have All the Data Gone? A Critical Reflection on Academic Digital Research in the Post-API Age. *Social Media + Society*, 7(1),

2056305121988929. <https://doi.org/10.1177/2056305121988929>

Venkatagiri, S. (2023, February 23). Researcher beware: Four red flags with the TikTok API's terms of service [Substack newsletter]. Technomoral.

<https://technomoral.substack.com/p/researcher-beware-four-red-flags>

Wassink, A. B., Gansen, C., & Bartholomew, I. (2022). Uneven success: Automatic speech recognition and ethnicity-related dialects. *Speech Commun.*, 140, 50–70.

<https://doi.org/10.1016/j.specom.2022.03.009>

X Developer Platform (2024). Non-commercial use of the X API. X.

<https://developer.x.com/en/developer-terms/commercial-terms>.

Yang, C. (2022). Bias in short-video recommender systems: User-centric evaluation on TikTok [Master's Paper, University of North Carolina].

https://cdr.lib.unc.edu/concern/masters_papers/v405sm32s

Zeitsoff, T. (2011). Using social media to measure conflict dynamics: An application to the 2008–2009 Gaza conflict. *Journal of Conflict Resolution*, 55(6), 938–969.

<https://doi.org/10.1177/0022002711408014>

Zhao, Y. (2024). TikTok and researcher positionality: Considering the methodological and ethical implications of an experimental digital ethnography. *International Journal of Qualitative Methods*, 23, 16094069231221374.

<https://doi.org/10.1177/16094069231221374>

Appendices (Chapter 1):

1. Selected TikTok Research Tools Terms of Service¹¹

ITEM NO.	SELECTED TERMS
II-4	"4. Subject to these Research Terms, TikTok hereby grants you a personal, non-exclusive, non-transferable, non-sublicensable, limited, revocable license during the approved Research solely to access and use the TikTok Research Tools for the sole purpose of undertaking the approved Research."
III-3C	"c. Publication of Research. Consistent with Open Science Principles, you agree to use reasonable efforts to publish any Research in Open Access journals or publications and/or as Open Access resources on other websites."
III-3E	"e. Data Refresh and Deletion. Where you are able to access or download Personal Data through the TikTok Research Tools, you agree to regularly refresh TikTok Research Data at least every thirty (30) days and delete data that is not available from the TikTok Research Tools at the time of each refresh. TikTok Research Data shall not be kept by you for longer than is necessary for Research approved as part of your application. You agree to provide TikTok with written certification of data deletion upon TikTok's request."
III-3F-II	"ii. You agree to provide TikTok with a copy of any publications pertaining to or containing the results and findings of the Research outputs, and any supporting information, at least seven (7) days before publication primarily to identify any user private Personal Data that may need to be removed prior to publication or disclosure."
III-3H	"h. License to TikTok. After you publish any research outputs, you agree that TikTok will have free and unlimited access to and use of your publication and research outputs. Such use may include, without limitation, the right to:
III-4-VII	"vii. not publish your Research outputs if your outputs can be linked back to a specific TikTok user or other individual;"
VI	"To the maximum extent permitted by applicable law, you agree to hold harmless and indemnify TikTok, its subsidiaries, affiliates, licensors, licensees, assigns and successors, and each of their officers, directors, employees, agents, advisors and shareholders from and against any third-party claim, suit or action including any liability, losses, damages (actual and/or consequential), expenses, litigation costs and reasonable attorney fees, of every kind and nature arising from or in any way related to your (a) actual or alleged breach of your obligations under these Research Terms; (b) your use of the TikTok Research Tools or data derived therefrom; or (c) the Research. TikTok shall use reasonable endeavours to provide you with written notice of any such claim, suit, or action."

2. Sampling procedure:

¹¹ <https://www.tiktok.com/legal/page/global/terms-of-service-research-api/en>

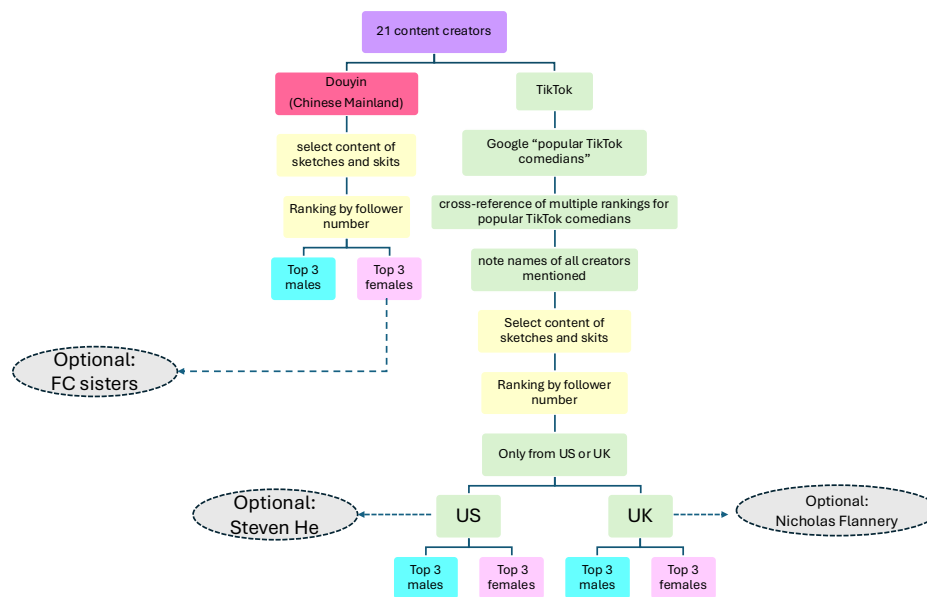


Figure 14 Sampling (TikTok and Douyin)

3. Endpoints for TikTok Research API

Category	Data returned	Query	Endpoint
Video	Videos configured to "Everyone" for who can watch, and for such videos, the total number of likes, total number of comments, voice-to-text, subtitles, the time of creation and video length.	Query Videos endpoint:	https://open.tiktokapis.com/v2/research/video/query/
Comments	comment text, and total number of likes, replies, and the time the comment was posted.	Query Video Comments:	https://open.tiktokapis.com/v2/research/video/comment/list/
Accounts	bios, profile pictures, liked videos, reposted videos, pinned videos, the total number of followers and the number of people who are followed by the account. The response also returns information on the followers and people followed by the account.	Query User Info:	https://open.tiktokapis.com/v2/research/user/info/
		Query User Liked Videos:	https://open.tiktokapis.com/v2/research/user/liked_videos/
		Query User Pinned Videos:	https://open.tiktokapis.com/v2/research/user/pinned_videos/
		Query User Followers:	https://open.tiktokapis.com/v2/research/user/followers/
		Query User Following:	https://open.tiktokapis.com/v2/research/user/following/

Other	The public playlists from accounts to get information on all videos that are part of a playlist.	Query User Reposted Videos:	https://open.tiktokapis.com/v2/research/user/reposted_videos/
		Query PlayList Info:	https://open.tiktokapis.com/v2/research/playlist/info/

4. Data Dictionary

	Variable Names	Annotation	Notes
Query Videos			
1	id	The unique identifier of the TikTok video. This is also called "item_id" or "video_id".	
2	create time	This is the time when the video was created.	
3	username	This is the username of the video creator.	
4	region code	A two digit code for the country where the video creator registered their account. Click here to learn more about the region codes returned.	
5	video description	This is the description of the video.	
6	music id	This is the music_id used in the video.	
7	like count	The total number of likes on a TikTok video, created by users by clicking the “Heart” icon.	
8	comment count	This is the total number of comments posted on a video.	
9	share count	The total number of times a TikTok video has been shared by clicking the "Share" button with the video.	
10	view count	This is the total number of views for a video on TikTok.	
11	effect_ids	The list of effects applied on the video.	such as filters
12	video id	This is a unique video ID for each video posted on TikTok. This is a number that can be used to reconstruct the URL link to access the video.	the same as "ID"
13	hashtags	The list of hashtags used in the video.	
14	hashtag_id	Returns the unique hashtag_ids for each hashtag.	
15	hashtag_description	Returns a description for a hashtag_name if one exists.	
16	video_mention_list	Returns the other tagged users in a video.	
17	video_label	Returns any labels applied to a video such as "election labels" (Ex: Get info on the U.S elections)	
18	playlist_id	The ID of the playlist that the video belongs to./The unique ID of the playlist.	
19	voice_to_text	Voice to text and subtitles (for videos that have voice to text features on, show the texts already generated), can be regarded as “transcription”	
20	is_stem_verified	Whether the video has been verified as being high quality STEM content.	
21	video_duration	The duration of the video in seconds.	
22	favorites_count	The number of favorites a video receives.	
Query Users			
1	following count	This is the number of people that a public user follows	
2	likes count	This is the total number of likes accumulated by the user	
3	video count	This is the total number of videos that the user has posted on their TikTok account	
4	bio description	This is the description in the bio of the user	If the user does not have a description, this will be returned blank
5	display name	This is the user’s profile name that is found under the username	
6	follower count	This is the total number of followers that follow the user	
7	avatar url	This is the URL of the user’s profile picture	
8	is verified	This returns the information on whether the user has been verified	All verified users will have "blue tick" next to their username
9	bio_url	The public URL in the user’s bio will be shared here	
Query Followers List			
1	display name	This is the profile name of the follower who follows the queried user	
2	username	This is the username of the follower who follows the queried user	
Query Following List			
1	Display Name	This is the profile name of the user who the queried user follows	
2	Username	This is the username of the user who the queried user follows	
Query Playlist Info			
1	playlist_id	The ID of the playlist that the video belongs to./The unique ID of the playlist.	
2	playlist_item_total	Provides the total number of items in a playlist	
3	playlist_last_updated	Provides info on when the playlist was last updated	
4	playlist_name	The name of the playlist	
5	playlist_video_ids	Provides a list of all video IDs in a playlist	
Source:	https://developers.tiktok.com/doc/research-api-codebook?enter_method=left_navigation		

5. Topic labels for TikTok videos

Topic Labels	Count of topic
accent	142
advertising	54
animal	70
behind the scenes	28
branding	73
childhood	220
driving	23
family	546
festival	7
food	137
friendship	67
history	18
home setting	104
interactive	5
intrusive thoughts	10
lockdown	8
mass media	321
monologue	68
other	15
personification	5
physical body	66
platform features	254
political	42
prank	18
relationship	459
social occasions	743
sports	162
stereotype	86
travel	15
user experience of technology	372
vlog	71
work	3
(blank)	
Grand Total	4212

6. A- Variable Quality: View Count (further tests)

When the two optional candidates were excluded, reducing the sample to eight creators, the pattern remained (see details in appendix xxx). API counts were again significantly lower than manual counts ($\beta = -0.185$, $p < 0.001$, 95% CI $[-0.248, -0.123]$), with no significant gender or nationality effects. Year effects were again substantial: view counts were higher in 2021 than 2020 ($\beta = 1.697$, $p < 0.001$, 95% CI $[1.594, 1.799]$) and 2022 ($\beta = 0.198$, $p < 0.001$, 95% CI $[0.109, 0.287]$) than 2021, while 2022 had markedly higher view counts than 2020 ($\beta = 1.697$, $p < 0.001$, 95% CI $[1.594, 1.799]$).

B- Variable Quality: Comment Count (further tests)

When the two optional candidates were excluded, reducing the sample to eight creators, the pattern remained. API counts were again significantly higher than manual counts ($\beta = 0.205$, $p < 0.001$, 95% CI $[0.136, 0.274]$), with no significant gender or nationality effects. Year effects were again pronounced: relative to 2020, comment counts were higher in 2021 ($\beta = 1.721$, $p < 0.001$, 95% CI $[1.608, 1.833]$) and 2022 ($\beta = 1.388$, $p < 0.001$, 95% CI $[1.280, 1.496]$). Using 2021 as the reference year confirmed these differences, with 2020 markedly lower ($\beta = -1.721$, $p < 0.001$, 95% CI $[-1.833, -1.608]$) and 2022 lower still ($\beta = -0.333$, $p < 0.001$, 95% CI $[-0.431, -0.235]$) than 2021.

7. Supplementary Tools for Multimodal Data

Website	Address	Data acquired	Input
SSTIK	https://ssstik.io/en-1	video and audio	video URL
tokdownload	https://tokdownload.com/	video and cover image	video URL
tokscript	https://script.tokaudit.io/	transcription	video URL
TikTok To Text	https://www.tiktottotext.com/	transcription	video URL or video
Happy Scribe	https://www.happyscribe.com/	transcription	video URL
otter.ai	https://otter.ai/	transcription	video
veed.io	https://www.veed.io/	transcription	video
Google Cloud Video Intelligence API	https://cloud.google.com/video-intelligence	text on the cover page	video
Mojibake decoder	https://www.linestarve.com/tools/mojibake/	Readable text	text

To be specific, linguistic analysis can be achieved. For example, speaking rate is available, where the number of words of transcriptions (voice_to_text) is divided by the number of minutes (converting “video_duration” from seconds to minutes). Researchers need to select the data without background music but only meaningful utterances. Relevant corpora can be established based on genres and creators, with video descriptions, transcriptions, and hashtag names. Thus, lexical complexity (Bottini, 2022), collocation and word frequency (Collins and Semino, 2024), and discourse analysis (Leedham et al., 2020) (Markowitz and Hancock, 2014) are achievable. Popular R packages are “quanteda”, “text2vec”, “tidytext”, “stringr”, and “spacyr” (Rul, 2019; Liu, 2024), along with traditional tool Praat by Paul Boersma and David Weenink, or free and more user-friendly one “#LancsBox X” developed and maintained by Lancaster University. Other options can be found in this directory: <https://corpus-analysis.com/>.

In social sciences, content and thematic analysis can be achieved by manual coding or with tools such as NVIVO or ATLAS.ti. Sentiment analysis of video descriptions and transcriptions is also feasible.

Alternatively, researchers can use the existing semantic tagging system (UCREL Semantic Analysis System, USAS) in WMatrix (Beigman Klebanov et al., 2007; Hou, 2019; Hayes and Poole, 2022). Semiotic analysis (mainly in qualitative studies) can show connotative and denotative meanings of TikTok videos (Markowitz and Hancock, 2014; Abril, 2023; Noor et al., 2024). One overlapping topic on TikTok between linguistics and social sciences is “Algospeak,” where creators use online “euphemism” to avoid content moderation (Steen et al., 2023; Klug et al., 2023; Felaco and Pelliccia, 2024). Profile information of creators mainly contributes to descriptive statistics, including follower base, total number of likes, and verification. Network analysis can be conducted to see the overlapping fanbase of different accounts or regarding a specific event, such as the 2010-2011 Australian Floods (Cheong and Cheong, 2011) and the 2016 Louisiana flood (Kim and Hastak, 2018)

In machine learning, there are ready-made models for speech emotion recognition, such as “SamLowe/roberta-base-go_emotions¹².” Social scientists can try the facial emotion recognition model¹³, which can show the range and change of five emotions considering limited texts from videos. Other studies involve NLP to monitor public opinion (Xu et al., 2023), LLM to Counter Algospeak (Fillies and Paschke, 2024), and computer vision to detect misinformation (Shang et al., 2021), all of which can be valuable references for social scientists. Specifically, results from big data facilitate semiotic analysis, topic categorisation and political communication.

12 Speech emotion recognition: SamLowe/roberta-base-go_emotions https://huggingface.co/SamLowe/roberta-base-go_emotions

13 ElenaRyumina/face_emotion_recognition https://huggingface.co/ElenaRyumina/face_emotion_recognition

Chapter 3: Sensory, Cognitive, and Social Verdicts: Multimodal Predictors of Likes, Comments, and Shares in TikTok Comedy

Abstract

Short-form video platforms now shape not only digital entertainment but broader cultural participation. TikTok, in particular, has created a highly competitive attention economy in which engagement signals determine content reach. Yet how multimodal features collectively drive content success on such platforms remains poorly understood. This study examines how visual composition, audio context, linguistic markers, platform affordances, and creator gender jointly predict engagement in TikTok comedy videos. We analysed 741 videos from 14 high-profile creators during the COVID-19 pandemic. Generalised Additive Mixed Models with creator-specific random intercepts were applied to model likes, comments, and shares as distinct outcomes. The findings show that these metrics are not interchangeable; they represent qualitatively distinct audience behaviours driven by different multimodal configurations. Likes function as sensory verdicts primarily shaped by visual composition, prosodic expressiveness in female creators, and video description length. Comments reflect cognitive verdicts driven by video duration, title sentiment, and gender-differentiated linguistic complexity. Shares embody social verdicts shaped by relationship-focused content, transcript sentiment, and gender-specific lexical properties. Video description length was the sole universal predictor across all three models. Gender consistently moderated multimodal effects, with female creators' engagement shaped more by prosodic and linguistic complexity and male creators' engagement depending more on lexical accessibility. These findings advance multimodal communication theory and provide practical guidance for content creators navigating algorithm-mediated platforms.

Keywords: TikTok, multimodal engagement, comedy, gender differences, social media metrics, algorithm-mediated communication

Introduction

The rapid rise of short-form video platforms has transformed cultural participation well beyond the domain of digital entertainment. TikTok has emerged as the most influential force in this shift. Boasting over 1.5 billion monthly active users by 2025 (Singh, 2025), the platform shapes societal discourse, influences consumer behaviour, and redefines creative expression. Unlike traditional social media platforms, which rely primarily on follower networks for content dissemination, TikTok prioritises engagement signals over social connections (Guinaudeau et al., 2022). This architectural distinction creates conditions for viral content emergence even from creators with modest followings.

These distinctive platform affordances create novel challenges for understanding engagement (Bhandari & Bimo, 2022; Vijay & Gekker, 2021). Short-form TikTok videos simultaneously deploy textual, visual, and audio elements within compressed temporal frames, typically ranging from 15 seconds to three minutes.

Within this highly competitive attention economy, creators must rapidly capture viewer interest whilst

navigating algorithmic systems, trending sounds, and genre conventions. Comedy is among the most prevalent categories on the platform (Shutsko, 2020). It exemplifies these dynamics through timing, audiovisual synchronisation, linguistic creativity, and cultural resonance.

Despite TikTok's growing influence, scholarly understanding of how multimodal features jointly shape audience engagement remains limited. Engagement is widely recognised as multidimensional, encompassing cognitive, emotional, and behavioural dimensions measurable through likes, shares, views, and comments (Agrawal, 2024; Chen et al., 2021; Cheng & Li, 2024; Shahbaznezhad et al., 2021; Vinerean & Opreana, 2021). Yet existing research has predominantly adopted fragmented, unimodal approaches, examining isolated features such as hashtags or visual aesthetics without considering their synergistic effects (Klug et al., 2021; Omar & Dequan, 2020; Schellewald, 2021). This is problematic, given that engagement emerges from the complex interplay between creator characteristics, visual composition, audio context, linguistic markers, and platform affordances (Xiao et al., 2025).

A related gap concerns the role of creator gender. Research demonstrates that male and female creators deploy multimodal features differently in audience-directed communication, spanning acoustic, linguistic, and gestural dimensions (Zhang et al., 2025). Audiences may respond differently to these gendered communication styles, yet gender has rarely been examined as a moderating factor in TikTok engagement research.

Genre specificity presents a further blind spot. Comedy content has received limited scholarly attention (Lockyer & De Benedictis, 2023), despite its engagement dominance on TikTok (Su et al., 2022) and its status as a globally significant art form (Collins, 2024). Comedic videos employ distinctive multimodal strategies, including timing mechanisms, audiovisual incongruity, linguistic wordplay, and cultural references (Chłopicki & Brzozowska, 2021; Su et al., 2022; Vásquez & Aslan, 2021). These strategies may operate through engagement pathways that differ from those documented in other content categories.

The present study addresses these research gaps through a multimodal analysis of TikTok comedy content from male and female creators during the COVID-19 pandemic. The pandemic provides a critical context for examining societal responses under collective stress, including patterns of isolation, mental health strain,

and misinformation circulation (Davidson et al., 2023). Comedy on TikTok, often rooted in shared social truths, offers a distinctive lens for interrogating cultural narratives (Berwick, 2024; Han & Kuipers, 2021). Two research questions guide the investigation. The primary question asks: how do visual composition, audio context, linguistic markers, and platform affordances jointly predict engagement in TikTok comedic videos? The subsidiary question asks: to what extent does creator gender moderate these multimodal effects? Together, these questions acknowledge that engagement emerges from synergistic multimodal interaction within TikTok's algorithm-driven ecosystem. They further recognise that gender may shape how audiences respond to these multimodal configurations.

Three hypotheses guide the present investigation. The first hypothesis, drawing on evidence from online video contexts (Xiao et al., 2025), posits that multimodal cues collectively contribute to TikTok engagement, though with differing weights. The second hypothesis posits that these effects will differ across engagement metrics. Likes represent low-effort responses, whilst comments and shares require deeper cognitive and social engagement, suggesting qualitatively distinct predictive profiles. The third hypothesis posits that gender will moderate multimodal effects, given established differences in multimodal production between male and female creators (Zhang et al., 2025).

The paper proceeds as follows. Section 2 reviews literature on factors contributing to TikTok engagement, whilst Section 3 outlines the methodology and Section 4 presents the results. Section 5 discusses implications, and Section 6 concludes.

Literature Review

TikTok engagement arises from an interplay of platform architecture, creator characteristics, and multimodal content features (see Table 1). Two contextual factors warrant brief mention before examining multimodal literature. First, TikTok's algorithm amplifies content based on engagement signals rather than follower networks. Early indicators such as likes and comments trigger wider distribution through feedback loops (Boeker & Urman, 2022; Chen, 2024; Guinaudeau et al., 2022; Klug, 2020). Second, creator attributes, including nationality, ethnicity, audience motivations, and personality traits, introduce systematic variation in how content is produced and received (Atayeva, 2024; Bhandari & Bimo, 2022; Katiaj et al., 2024; Li et

al., 2026; Meng & Leung, 2021; Wahid et al., 2022). These factors condition the contexts in which multimodal features operate. Yet the central concern of this review is the content itself.

2.1 Multimodal Content Characteristics

TikTok videos integrate visual, auditory, and linguistic information simultaneously within compressed temporal frames. Multimodal communication treats meaning as arising from the co-deployment of these semiotic modes rather than from any single channel in isolation (Chen, 2024; Wang, 2021). Prior research has predominantly examined visual aesthetics, audio elements, or linguistic markers separately, treating each as independent of the others. This fragmentation limits understanding of how cross-modal interactions shape audience response.

Content authenticity is central to engagement on TikTok. Audiences are more likely to interact with material they perceive as genuine rather than overly polished (Agrawal, 2024). Emotionally charged content, including material evoking humour, surprise, or inspiration, consistently generates higher engagement (Atayeva, 2024; Barta et al., 2023; Eide & Jonassen, 2024). Yang et al. (2024) further demonstrate that influencers expressing both humour and frustration create emotional landscapes that encourage active participation. These affective dynamics operate across modalities, shaping how visual, audio, and linguistic elements collectively signal authenticity and emotional investment.

Visual composition

Visual composition represents one dimension through which emotional content is conveyed. Facial expressions and gestural behaviour contribute to perceived warmth and interactional involvement in video communication (Nagels et al., 2023; Zhang et al., 2025). These nonverbal cues directly signal emotional states and relational intent. However, they remain difficult to operationalise at scale. Gestural analysis requires intensive manual coding. Beyond these emotion-laden features, compositional properties also shape engagement. The rule of thirds is a compositional guideline whereby key visual elements are positioned along imaginary gridlines dividing the frame into thirds. This framing principle is associated with higher perceived aesthetic quality and viewer attention across platforms (Kostyk & Huhmann, 2021; Shen & Pritchard, 2022; Simonetti & Bigne, 2022). Brightness and contrast provide additional perceptual cues that

direct viewer attention (Kostyk & Huhmann, 2021; Shen & Pritchard, 2022). Wang (2021) shows that broader visual aesthetics shape audience interpretation. Platform-native features such as filters and on-screen text (including subtitles) shape the aesthetic register of content and signal creator familiarity with platform conventions (Ahlse et al., 2020; Ngilangil, 2022). Li et al. (2021) demonstrate that TikTok videos using visual aids such as subtitles and clear messaging enhance comprehension and sharing behaviours. Nevertheless, the extent to which these platform-specific visual affordances predict engagement within an integrated multimodal framework and how they may be moderated by creator gender remains untested.

Audio and prosodic features

Prosodic properties of speech represent a theoretically grounded but empirically underexplored set of engagement predictors. Audio loudness aligns with arousal-attention-memory models, in that higher loudness increases physiological arousal and enhances attentional focus and recall (Mather & Sutherland, 2011). Recent computational media research confirms the utility of loudness measures in capturing the perceptual dynamics of audiovisual content (Lee et al., 2024). Beyond loudness, original audio elements, such as soundtracks and trending music, enhance narrative coherence and emotional resonance (Wahid et al., 2022). Original audio content tends to outperform repurposed material in generating engagement, owing to stronger audio-visual alignment within a given video (Wahid et al., 2022). However, other prosodic cues, such as pitch, speech rate, and waveform properties, by contrast, are seldom examined as explicit predictors in social media contexts. These cues are found to reflect speakers' emotion (references here), but their specific contributions to audience engagement remain largely unknown.

Table 5 Literature on TikTok/Douyin Engagement Metrics

Factors influencing TikTok Engagement Metrics		Sample Size	Method	Data Type		
<u>Content Characteristics</u>				Textual	Audio	Visual
sentiment; content characteristics, and the role of influencers.	(Agrawal, 2024)	2000 videos	sentiment analysis; <i>go_emotions</i> model and NLTK VADER.	Yes		
humour, relatability, and personal experience	(Barta et al., 2023)	217 TikTok followers	survey	Yes		
emotional resonance (humour, surprise, or inspiration); the role of influencers, and the strategic timing and relevance of content	(Eide and Jonassen, 2024)	11 interviewees	Qualitative: interviews	Yes		
express both humour and frustration	(Yang et al., 2024)	223 videos	ANOVA, Chi-square, Dickey – Fuller test	Yes	Yes	
visual aesthetics	(Wang, 2021)	5 videos	Qualitative: semiotic storyboard			Yes
caption (video description) sentiment and length	(Chen, 2024)	110 college students	Survey	Yes		
Emotional content; rational content; original audio, including trending music and custom soundtracks	(Wahid et al., 2022)	100 videos	mixed-effect model	Yes	Yes	
human characters, use of subtitles, and the integration of clear risk messaging	(Li et al., 2021)	331 videos	Correlation and ANOVA analyses	Yes	Yes	Yes
strategic integration of creative tools such as filters, catchy background music, and product demonstrations	(Agabo, 2024)	210 participants	qualitative content analysis, discourse analysis and key informant interviews	Yes		
user-friendly features: interactive challenge	(Ahlse et al., 2020)	16 interviewees	Qualitative: interviews, grounded theory	Yes		
user-friendly features: duets, stitches	(Ngilangil, 2022)	72 participants	Survey	Yes		
Sentiment (negative); second-person pronouns	(Cheng & Li, 2024)	101,292 videos	Sentiment Analysis; facial analysis; linear regression models	Yes		
Video length, titles’ length and sentiment (positive), dialogic loop, and content type	(Chen et al., 2021)	355 videos	content analysis, negative binomial regression	Yes		
multimodal information features (sentiment score; fluctuant frequency, pitch, audio diversity; brightness, contrast, saturation)	(Xiao et al., 2025)	4292 veridical short sales videos	multiple regression analysis (MRA) and fuzzy set qualitative comparative analysis (fsQCA)	Yes	Yes	Yes
<u>Algorithmic Recommendation</u>						
TikTok’s ‘For You’ page prioritizes personalized content based on various user interaction signals	(Guinaudeau et al., 2022)	2,026,506 videos from 11,597 political accounts	Correlation analyses	Yes		
user engagement metrics shape the recommendation algorithm’s candidate selection and subsequent ranking	(Boeker & Urman, 2022)	30,436 videos	sock-puppet audit methodology with a custom algorithm	Yes		
user characteristics (e.g. language settings and geographic location) are also factored into the algorithm’s content distribution	(Zannettou et al., 2024)	9.2M TikTok video recommendations from 347 TikTok users	data donation; correlation analyses	Yes		
<u>User Motivations and Psychological Drivers</u>						
Creator attributes, such as age, gender, and nationality, alongside cultural contexts	(Bhandari & Bimo, 2022)	14 interviewees	Qualitative: interviews	Yes		
Influencers’ follower bases, close-up and medium-shot, the lifespan of the video, the presence of text, and the point of view.	(Ling et al., 2022)	400 videos	mixed-method	Yes		Yes

spontaneous viral content, personality traits	(Meng and Leung, 2021)	526 participants	questionnaire and correlation analyses	Yes
visually appealing, short, authentic, emotionally impactful, and often humorous content	(Atayeva, 2024)	11 interviewees	Qualitative: interviews	Yes
Motivational factors (e.g. social interaction, escapism, self-expression, and archiving behaviour)	(Katiaj et al., 2024)	202 videos from 17 accounts	ANOVA analyses	Yes

Gender differences in prosodic behaviour are well established across languages and contexts. Women consistently exhibit higher mean pitch, wider pitch range, and greater pitch variability than men (Demina, 2021; Leyns et al., 2024). These patterns reflect both anatomical differences in vocal fold structure and sociocultural expectations around expressive intonation (Leyns et al., 2024; Neer et al., 2025). Female speakers show more dynamic pitch patterns in both read and spontaneous speech (Leyns et al., 2024; Meer & Fuchs, 2022). They tend to use more frequent pitch shifts to maintain social harmony, whilst men more often adopt monotone intonation (Demina, 2021). Evidence on speech rate is less consistent across studies and languages (Abdel-Hamid et al., 2020; Angelopoulou et al., 2020; Damhoureyeh et al., 2020; Mönnich et al., 2023), with findings varying by context rather than by gender alone. To the authors' knowledge, however, these gendered prosodic differences have not systematically been examined as explicit predictors of engagement in TikTok research.

Linguistic features

Linguistic features constitute a third dimension through which emotional content is conveyed. Cheng and Li (2024) find that sentiment and second-person pronouns predict engagement with TikTok news videos, with direct address enhancing interpersonal involvement and perceived relevance. Chen et al. (2021) identify video length, title characteristics, dialogic loops, and content type as significant predictors. Storytelling grounded in authenticity and relatability enhances interaction and dissemination (Heikkonen, 2024; Katiaj et al., 2024).

Finally, type-token ratio and word frequency are established indicators of communicative complexity and accessibility in language research (Asaadi et al., 2024; Carrasco-Farré, 2022). Neither has been applied to TikTok comedy as an explicit predictor of engagement. Their exclusion from existing TikTok studies reflects a broader tendency to subsume such measures within larger natural language processing pipelines rather than examining them independently. The present study includes these measures directly, alongside sentiment, pronoun usage, and title characteristics, within a unified multimodal model. Crucially, it treats likes, comments, and shares as qualitatively distinct outcomes rather than collapsing them into a single metric.

Gender differences in linguistic and affective behaviour are equally well documented. Women lexically express more positive emotions and use more affective, supportive language in online communication than men (Al-Rawi et al., 2021; Sun et al., 2020). Women are also more sensitive to emotional cues and adapt their sentiment expression accordingly, whilst men tend towards more rational, less emotionally inflected responses, particularly in technical or information-seeking contexts (Al-Rawi et al., 2021; Feng et al., 2023). Zhang et al. (2025) extend these findings to broadcast contexts, showing that male and female presenters produce distinct multimodal profiles across acoustic, linguistic, and gestural dimensions, with differences arising from a combination of biological and sociocultural factors. These gendered communication patterns suggest that male and female TikTok creators may deploy multimodal features differently, and that audiences may respond in ways that current models have yet to capture. However, how creator gender interacts with acoustic, linguistic, and affective features in predicting engagement is unknown.

2.2 Research Gaps and Study Positioning

The most persistent limitation in understanding TikTok engagement is the predominance of unimodal analyses (Chen, 2024; Wang, 2021). Studies that examine visual, audio, or linguistic features in isolation cannot fully capture the synergistic effects that emerge when these channels operate together. A theoretically coherent account of engagement requires models that treat multimodal features as an integrated rather than additive system. This will not only be more ecological in studying platform engagement but provides a better picture of multimodal communication and its impact on audience.

Xiao et al. (2025) represent the closest existing precedent, using multiple regression and fuzzy-set qualitative comparative analysis to examine multimodal features in short sales videos. Their work is valuable while focusing on a commercial focus without separating gender that constrains applicability to gender-specific entertainment genres such as comedy. Two methodological limitations further restrict generalisability. First, their modelling of nonlinearity relied on pre-specified quadratic terms, which may not adequately reflect more complex curvilinear relationships. Second, the model treated videos as independent observations, disregarding the non-independence of content produced by the same creator. So far, few

studies have additionally combined creator-level variables, specifically creator gender and nationality, with visual, audio, and linguistic predictors within unified engagement models.

Additionally, comedy content remains underexamined relative to its prominence on TikTok. Comedy represents a dominant engagement category on the platform (Shutsko, 2020; Su et al., 2022), yet most engagement research centres on informational or educational content (Li et al., 2021; Ming et al., 2024), leaving comedic dynamics poorly understood. Comedy relies on multimodal strategies such as timing mechanisms, audiovisual incongruity, and linguistic wordplay (Vásquez & Aslan, 2021). These may operate through different engagement pathways than those documented in other content

A further theoretical concern is the treatment of engagement metrics as interchangeable proxies. Likes, comments, and shares reflect qualitatively different degrees of cognitive and social investment. Most existing studies examine a single metric (e.g., Shi et al., 2023) or collapse multiple metrics into a composite (e.g., Lu et al., 2023; Wang, 2025), which obscure the possibility that different multimodal features drive different audience behaviours.

To address these gaps, the present study analyses 741 TikTok comedy videos to predict engagement across three distinct metrics: likes, comments, and shares. Two objectives guide the investigation. First, it examines how visual, audio, and linguistic features jointly influence each engagement metric. Second, it tests whether creator gender moderates the relationship between multimodal cues and engagement outcomes. Creator-specific random intercepts are incorporated to account for the non-independence of videos within creators, and nonlinear models are employed to detect complex relationships without imposing pre-specified functional forms. Together, these choices provide the kind of integrated, genre-sensitive analysis that current literature has yet to offer. The findings offer theoretical insights into digital comedy whilst providing practical guidance for content creators.

Methodology

3.1 Sample and Data Collection

Sample Composition

The study examines TikTok comedy content from 1 January 2020 to 31 December 2022, a period spanning the COVID-19 pandemic and its aftermath. Admittedly, the data are not experimentally induced conditions, but the observational design ensures that engagement patterns reflect authentic platform dynamics, which is also ecological valid (Bai & Gu, 2026; Davidson et al., 2023).

The sample comprised 14 high-profile comedy creators active on TikTok across the study period. Creators were identified through Google search visibility and viral content indicators, covering influential comedic voices from both the United Kingdom and the United States. The sample included six UK-based creators (three male, three female) and six US-based creators (three male, three female), plus two accounts designated as optional. The account @steven_he could not be unambiguously classified by nationality, identifying as Chinese-Irish whilst residing in the United States during the study period. The account @nicholas_flannery began uploading only in 2021, resulting in incomplete temporal coverage. Detailed sampling procedures are provided in Appendix 1.

Inclusion Criteria and Final Dataset

Videos were selected on the basis of two criteria: availability of transcription data and inclusion within each creator's top 10 per cent most-viewed content during the study period. Transcription data were obtained from the TikTok Research API or third-party speech-to-text services. Videos featuring purely visual humour without verbal content were excluded. The 14 creators published 6,373 videos collectively, of which 788 possessed complete transcriptions and met both inclusion criteria.

A further 47 videos proved unavailable during data collection in September 2024, due to age-restriction policies implemented by UK authorities on 25 July 2025 (UK Annex | TikTok, 2025), rendering them inaccessible via public TikTok webpages. The final dataset comprised 741 videos with complete multimodal feature extraction. Appendix 2 provides a detailed breakdown of transcription sources and video counts by creator.

Data Sources

Data were obtained from multiple complementary sources (Appendix 2) to ensure comprehensive coverage of creator content and engagement metrics. Metadata and performance metrics were acquired through two primary channels. The first dataset was obtained in March 2023 from Analisa.io, a commercial social media analytics service (Lachief, 2023). This service provided comprehensive metadata for creator profiles and video performance metrics. Access to several creator profiles was subsequently lost following the service's closure. The second dataset was collected in April 2024 via the official TikTok Research API, accessed using R packages including httr and jsonlite (Pascual, 2020). Data were filtered by publication date (2020–2022), geographic region (United Kingdom and United States), and key content metrics detailed in Appendix 2. Transcripts were obtained from two complementary sources. The TikTok Research API, accessed using R packages httr and jsonlite (Pascual, 2020), provided transcriptions for a subset of videos. However, owing to the API's unstable performance (Bai & Gu, 2026), additional transcripts were sourced from third-party services (happyscribe, otter.ai). Manual verification occurred between April and July 2023, identifying missing videos between website metadata and purchased or API-sourced data. This process ensured comprehensive coverage of each creator's output during the study period. Video files were not provided natively by either source. Consequently, videos were obtained using third-party downloading services¹⁴ and converted from MP4 to WAV format using the av package, enabling systematic audio feature extraction. This study has been ethically approved by the university (ETH2526-0311).

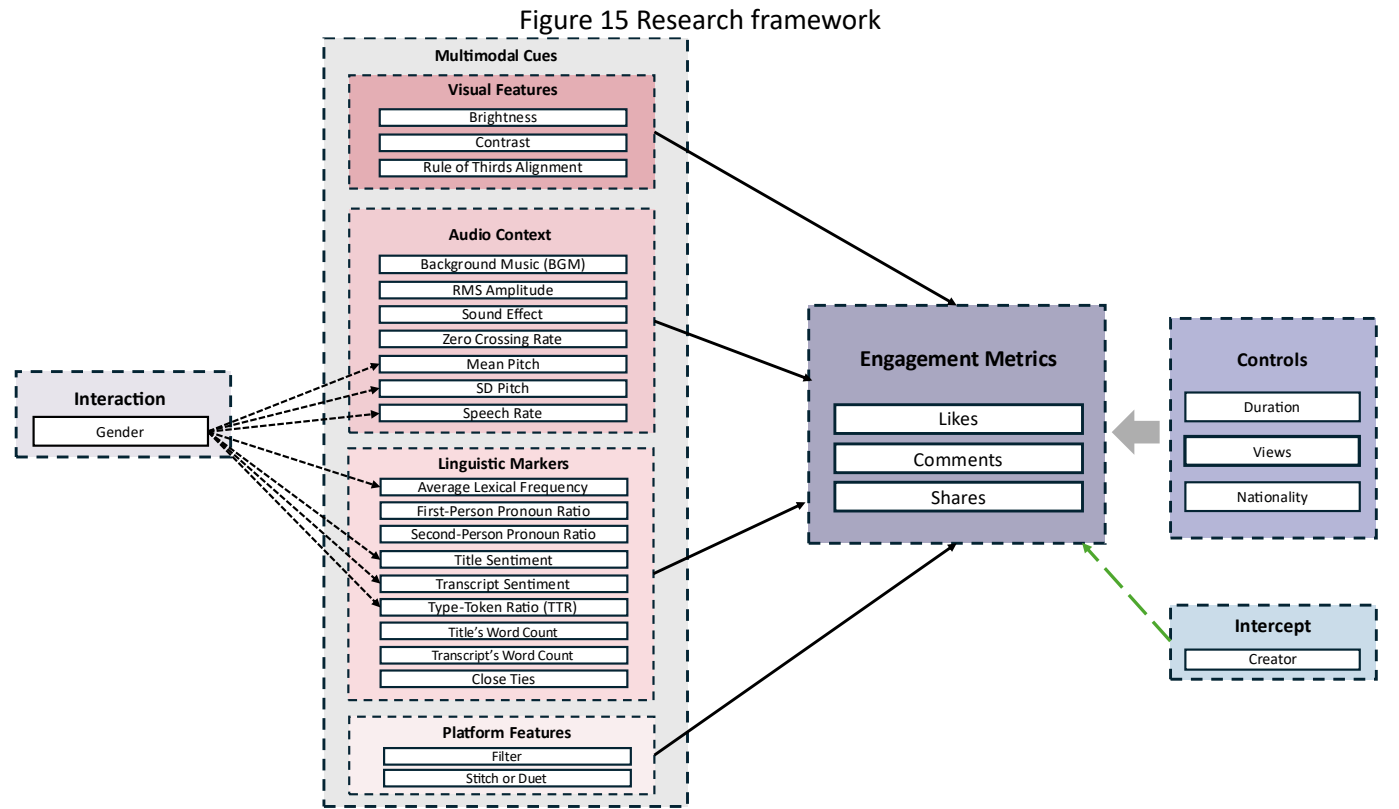
3.2 Measurement

Measurement Procedure

Variables were measured across five dimensions, as illustrated in Figure 1: visual composition features, audio context features, linguistic markers, platform affordance features, and creator-level demographic variables. Visual and audio features were extracted computationally from downloaded video files. Linguistic

¹⁴ tokdownload.com, videototext.io/tiktok-downloader

features were derived from transcriptions. Platform affordance features were hand-coded through direct observation by the researcher. Creator-level features (gender and nationality) were determined from publicly available profile information. Detailed variable information is provided in Table 2.



Note: the visualisation is inspired by Xiao et al. (2025)

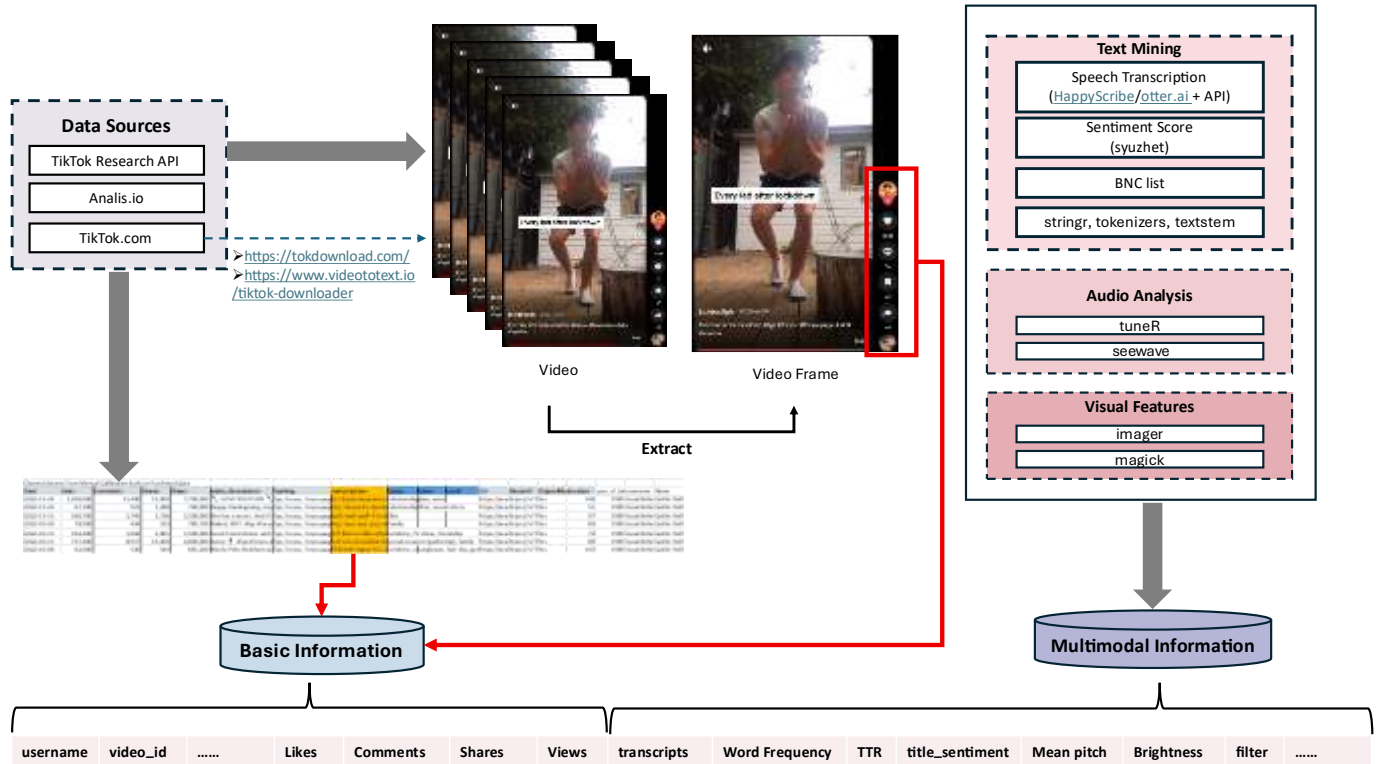
Visual features were extracted frame by frame using the `imager` and `magick` packages in R (Barthelme, 2015; Ooms, 2016; Xiao et al., 2025). Brightness was computed as the mean luminance of normalised greyscale pixel values across sampled frames, averaged over the full video (Watier, 2023). It is a dimensionless value between 0 and 1, where 0 represents complete darkness and 1 represents maximum light intensity (Bilozerskyi et al., 2023). Contrast was computed as the standard deviation of normalised greyscale pixel intensity values across sampled frames, averaged over the full video frame (Kostyk & Huhmann, 2021). It is likewise a dimensionless value between 0 and 1, where higher values indicate greater variation in luminance across the (Schwenzow et al., 2021). The rule-of-thirds score was computed as the mean per-pixel Sobel edge magnitude along the four rule-of-thirds gridlines, normalised by total frame size and averaged across frames sampled at two frames per second (Jin et al., 2022). It is a dimensionless ratio

where higher values indicate greater edge energy along compositional gridlines, reflecting more aesthetically balanced framing (Zhang et al., 2023).

Audio features were derived using the **tuner** and **seewave** packages (Ligges et al., 2004; Sueur et al., 2006). Loudness was measured as the root mean square (RMS) of the raw PCM waveform, computed from the integer sample values of the mono audio channel, with larger values indicating greater overall sound energy (Er, 2020; Hove et al., 2019). Zero-crossing rate reflects the number of times per second the audio waveform transitions between positive and negative values, measured in Hz, serving as an indicator of audio frequency complexity and rhythmic texture (Er, 2020; Rehman et al., 2021). Mean pitch and pitch standard deviation were extracted in Hertz using the **fund()** function from **seewave** alongside speech rate, measured in syllables per second (Sueur et al., 2006).

Linguistic features were computed from transcriptions using R packages **stringr**, **tokenizers**, and **textstem** for text mining operations (Mullen, 2016; Rinker, 2017; Wickham, 2009). BNC word frequency was calculated by comparing transcriptions against the British National Corpus frequency list (Leech et al., 2014). Sentiment analysis employed the **syuzhet** package, yielding numerical scores where positive values indicate positive sentiment (Jockers, 2015). First-person pronoun ratios (I, me, my, mine, us, our, ours, myself) and second-person pronoun ratios (you, your, yours, yourself, yourselves) were standardised by dividing respective pronoun counts by transcript word count (Cheng & Li, 2024; Chung et al., 2023). All binary platform affordance variables, including background music, sound effects, duets or stitches, and filters, were manually coded by the researcher from direct observation, ensuring accuracy for platform-specific affordances not reliably captured through automated extraction.

Figure 16 Procedure of data collection and processing



Note: the diagram is inspired by Xiao et al. (2025)

Variable Transformation

Transformations were applied to ensure variables were on comparable scales and that model assumptions were met. Preliminary examination revealed substantial positive skewness in engagement metrics and several predictor variables, necessitating logarithmic transformations to approximate normal distributions and satisfy regression assumptions. Engagement outcomes (likes, views, comments, and shares) were log-transformed using $\log(x + 1)$ to reduce skewness and stabilise variance (Gelman & Hill, 2006). Duration and count variables were log-transformed and then standardised. Proportional measures were also log-transformed where skewness was pronounced, then standardised.

For prosodic cues, speech rate was log-transformed, and raw pitch values were converted from Hertz into semitones relative to a 50 Hz reference using the formula (Shi, Gu, and Vigliocco, 2022):

$$ST = 12 \times \log_2\left(\frac{\text{Hz}}{50}\right),$$

Both semitone measures were subsequently standardised (z-scored) prior to analysis. Negatively skewed variables, including average lexical frequency and brightness, were squared before standardisation (Tabachnick & Fidell, 2019). Sentiment scores were centred and scaled only, as they are bounded around zero and unsuitable for log transformation (Smithson & Verkuilen, 2006; Tabachnick & Fidell, 2019).

Categorical variables, including nationality, gender, background music, filters, sound effects, stitch or duet use, and close ties, were converted to factors. This transformation strategy ensured that all predictors contributed appropriately to the statistical models whilst maintaining interpretability of effect sizes.

3.3 Data Analyses

Three engagement metrics served as outcome variables: like count, comment count, and share count. These metrics reflect qualitatively distinct forms of audience response. Likes indicate passive approval, comments signal active discursive participation, and shares represent content amplification through network diffusion (Habibi & Salim, 2021). Each outcome was log-transformed to address positive skewness and modelled as a function of multimodal features alongside appropriate controls.

Preliminary diagnostics assessed multicollinearity amongst predictor variables. Correlation analyses confirmed that predictors across modalities exhibited generally low to moderate associations (see Appendix 3), reducing concerns about severe multicollinearity in the regression models. Variance inflation factors (VIF) were computed for all continuous predictors, with all values falling below 5 (Quah et al., 2024), indicating acceptable levels of multicollinearity. One exception was noted: transcript word count exhibited substantial collinearity with video duration ($r = 0.72$, $p < .001$). This association reflects the inherent temporal constraint that longer videos accommodate more spoken content. Transcript word count was therefore removed from subsequent analyses, with duration retained as the more fundamental temporal measure.

Independent variables spanned five dimensions (Figure 1). (1) Creator-level features captured gender and nationality, demographic characteristics potentially influencing content style and audience reception. (2) Video-level visual features encompassed duration in seconds, rule-of-thirds score, contrast, and brightness, quantifying compositional aesthetics and technical characteristics. (3) Audio context included loudness (root mean square amplitude), zero-crossing rate, speech rate, mean pitch, pitch standard deviation, background music (binary), and sound effects (binary), measuring acoustic properties and sound design choices. (4) Linguistic markers comprised type-token ratio, BNC word frequency, video description length, transcript word count, video description sentiment, transcription sentiment, first-person pronoun ratio, second-person

pronoun ratio, and mean length of utterance in words, assessing lexical diversity, word frequency, sentiment valence, and interpersonal address patterns. (5) Platform features captured duets and stitches usage alongside filter usage, representing TikTok-specific affordances enabling collaborative content creation and visual enhancement. View count was included as a control variable.

Gender interaction terms were specified for eight theoretically motivated continuous predictors: speech rate, mean pitch, pitch standard deviation, type-token ratio, average lexical frequency, transcript sentiment, and video description sentiment. These interaction terms test whether the relationship between each of these multimodal cues and engagement differs systematically by creator gender. The remaining continuous predictors, including RMS amplitude, zero-crossing rate, duration, rule of thirds, brightness, contrast, first-person pronoun ratio, second-person pronoun ratio, and video description length, were modelled as main effects without gender interactions.

The study employed Generalised Additive Mixed Models (GAMMs) with creator-specific random intercepts to model log-transformed engagement outcomes. Figure 1 presents the overall analytical framework.

GAMMs extend traditional regression by incorporating smooth, nonlinear functions for continuous predictors whilst accommodating hierarchical data structures through random effects (Wood, 2017). The estimated degrees of freedom (edf) reported for each smooth term reflect the complexity of the fitted function: an edf of 1 indicates a linear effect, whilst values above 1 indicate progressively more complex curvature. Parametric terms, including categorical predictors and control variables, are reported with their regression coefficients (β) and p-values. Smooth terms are reported with their estimated degrees of freedom (edf) and approximate significance values, as these terms do not yield standard regression coefficients regardless of whether their estimated relationship is linear ($\text{edf} = 1$) or curvilinear ($\text{edf} > 1$). This approach allows detection of nonlinear relationships without requiring a priori specification of functional forms, accommodates the nested structure of videos within creators, and controls for creator-level heterogeneity. Each creator contributed between 15 and 113 videos, with a mean of 53 per creator, providing a within-creator structure that yields stronger estimation grounds than cross-sectional designs.

Models were estimated using restricted maximum likelihood (REML) (Shen & Lu, 2025) via the function in R's `gamm4` package (Wood & Scheipl, 2009). The general model form was:

$$\begin{aligned} \log(\text{Engagement}) &= \beta_0 + \beta_1 \log(\text{Views or Likes}) + f_1(\text{visual features}) \\ &+ f_2(\text{remaining audio features}) + f_3(\text{selected audio features, by} \\ &= \text{Gender}) + f_4(\text{remaining linguistic features}) \\ &+ f_5(\text{selected linguistic features, by} \\ &= \text{Gender}) + \beta_{\text{categorical}}(\text{nationality} + \text{filter} + \text{background music} \\ &+ \text{sound effects} + \text{stitch or duet} + \text{close ties}) + u_{\text{creator}} + \varepsilon \end{aligned}$$

Here f_1 represents smooth main-effect functions for all visual features. f_2 and f_4 represent smooth main-effect functions for audio and linguistic features respectively that were not specified with gender interactions. f_3 and f_5 represent smooth functions estimated separately by creator gender for the interacted predictors: speech rate, mean pitch, and pitch SD (audio); and type-token ratio, average lexical frequency, transcript sentiment, and video description sentiment. Each continuous predictor enters the model in only one of these two forms, either as a main-effect smooth or as a by-gender smooth, not both. u_{creator} captures creator-specific random intercepts, and ε represents residual error. All smooth terms used penalised regression splines ($k = 5$ basis functions).

Two nested specifications were estimated for each outcome: a comprehensive model incorporating all theoretically motivated predictors, and a parsimonious model retaining only statistically significant terms. Both specifications included random intercepts for individual creators, addressing within-creator correlation. Basis dimensions were increased for smooth terms approaching their limits, as identified through `gam.check` diagnostics. Non-significant terms were systematically removed to yield simplified models with improved parsimony. Final models retained only significant predictors ($p < 0.05$) or those approaching significance ($p < 0.10$) where theoretically justified. Model comparison employed AIC, with lower values indicating better fit.

Separate models were fitted for like count, comment count, and share count to examine whether multimodal features differentially predict distinct forms of audience engagement. The three models were pre-specified and treated as independent analyses addressing distinct outcome variables; no correction for multiple comparisons across models was therefore applied (Bender & Lange, 2001). For the likes model, log-transformed views served as the primary control. For the comments and shares models, log-transformed

likes served as the primary control, as these behaviours represent deeper forms of engagement beyond passive viewing. The random intercept structure partitions variance into between-creator and within-creator components, providing more accurate estimates of multimodal feature effects whilst controlling for creator heterogeneity and the non-independence of videos from the same creator.

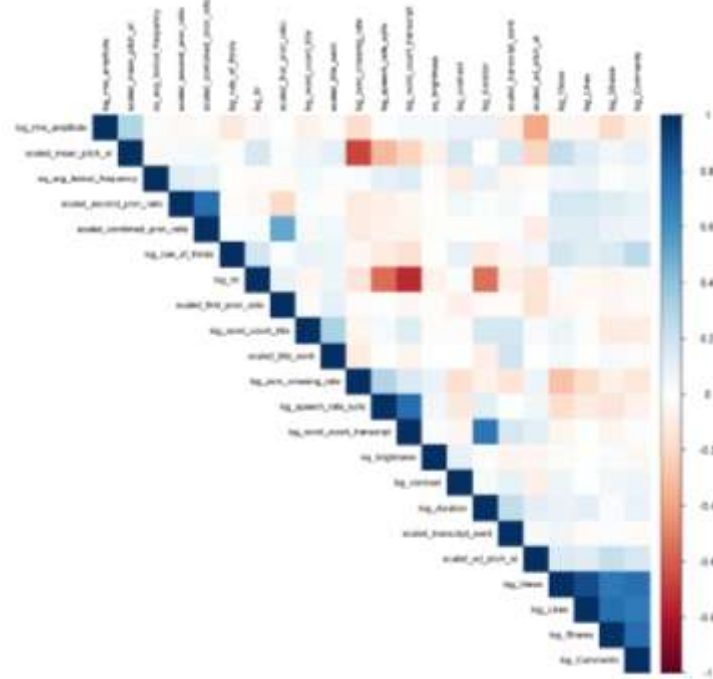
Results

4.1 Descriptive Statistics

Descriptive statistics for all variables are presented in Table 2. The engagement metrics and multimodal predictors demonstrated notable variability across the corpus, necessitating the transformations detailed in Section 3.2. A correlation analysis was conducted across text, audio, and visual modalities, with results displayed in Figure 3.

Strong positive associations were observed amongst the engagement metrics themselves. Log-transformed views and likes correlated at $r = 0.88$ ($df = 739$, $p < .001$), whilst shares and comments showed $r = 0.77$ ($df = 739$, $p < .001$). By contrast, most audiovisual and linguistic predictors exhibited weak to moderate correlations with engagement outcomes. Mean pitch and pitch variability related modestly to views and shares (r ranging from 0.06 to 0.24, all $p < .05$). Lexical diversity (log TTR) correlated negatively with duration ($r = -0.54$, $df = 739$, $p < .001$), reflecting the time constraints inherent to short-form video. Speech rate showed a weak positive association with duration ($r = 0.12$, $df = 739$, $p < .05$), indicating that faster-speaking creators tend to produce longer videos. Notably, TTR and speech rate were strongly negatively correlated with each other ($r = -0.56$, $df = 739$, $p < .001$), suggesting that faster speech is associated with lower lexical diversity. Visual composition features demonstrated negligible correlations with engagement. An exception was rule-of-thirds alignment, which showed a small positive association with comments ($r = 0.26$, $df = 739$, $p < .001$). Additional bivariate relationships are reported in Appendix 4.

Figure 17 Correlation Matrix



Note: Correlations were generally low to moderate ($|r| < .55$), indicating limited multicollinearity.

Whilst engagement measures clustered tightly together, the multimodal predictors contributed relatively independent sources of variance. This pattern reduces concerns about severe multicollinearity in the subsequent regression models.

Table 6 Descriptive statistics of variables

Variables	N	Mean	Std	Min	Max	Skewness	Category	Descriptions
Comments	788	90676.02	830359.10	0.00	12500000.00	11.13	Engagement, DV	The count of comments of a TikTok video
Shares	788	17184.84	59246.72	4.00	1300000.00	14.20	Engagement, DV	The count of shares of a TikTok video
Likes	788	601016.29	1008094.24	180.00	8500000.00	3.65	Engagement, DV and IV	The count of likes of a TikTok video
Views	788	5602318.39	12158358.00	4734.00	163900000.00	6.48	Engagement, IV	The count of videos of a TikTok video
duration	788	49.38	33.85	4.00	351.00	3.40	Video	how long the video is in seconds
nationality	788	1.51	0.58	1.00	3.00	0.65	Creator-Level Features	
gender	788	1.67	0.47	1.00	2.00	-0.71	Creator-Level Features	
brightness	741	0.48	0.12	0.00	0.86	-0.81	Video	the overall lightness or luminance of the frames
contrast	741	0.23	0.04	0.11	0.35	0.06	Video	the difference in luminance (light intensity) between the darkest and brightest parts of the frames.
rule.of.thirds_alignment	741	0.000487	0.000383	0.000073	0.005561	4.37	Video	Frames aligned with the rule of thirds are often perceived as more professional and visually engaging. A higher rule-of-thirds score indicates more aesthetically pleasing composition, whilst scores closer to zero suggest centred subjects or images with few distinct edges.
bgm	788	1.02	0.13	1.00	2.00	7.58	Audio	binary: whether the video uses background music or not (No=1, Yes=2)
rms_amplitude	741	3,725.08	2,016.44	136.92	13,164.26	1.44	Audio	Measuring loudness of speech or music in raw PCM amplitude units
sound_effect	788	1.02	0.13	1.00	2.00	7.58	Audio	binary: whether the video uses sound effects or not (No=1, Yes=2)
speech_rate_sylls	788	2.29	1.09	0.07	6.57	0.07	Audio	syllables per second (syll/s)
zero_crossing_rate	741	0.06	0.02	0.02	0.16	0.6	Audio	the frequency with which an audio signal transitions
mean_pitch	741	7.20	7.03	1.21	14.52	0.42	Audio	The average pitch of the given audio file in Hz
sd_pitch	741	4.88	4.94	1.23	7.05	-0.74	Audio	Standard deviation of pitch values of the given audio file: a small SD means the voice is relatively stable in pitch in Hz
avg_lexical_frequency	788	6.49	0.18	5.09	7.41	-1.07	Linguistic markers	The average number of times a given word in one transcript appears in BNC frequency list
close_ties	788	1.25	0.43	1.00	2.00	1.15	Linguistic markers	binary: Yes-videos concerning family, friendship and relationship (Chung et al., 2023); No-the rest of the topics (No=1, Yes=2)
first_pron_ratio	788	0.04	0.03	0.00	0.33	2.26	Linguistic markers	number of first-person pronouns (Chung et al., 2023) divided by word count of a given transcription
secon_pron_ratio	788	0.05	0.04	0.00	0.38	1.72	Linguistic markers	number of second-person pronouns (Cheng & Li, 2024) divided by word count of a given transcription
title_senti	788	0.15	0.75	-2.00	4.65	1.06	Linguistic markers	sentiment scores of a video description: positive value indicates positive sentiment and vice versa
transcript_senti	788	0.90	2.71	-8.30	14.10	0.5	Linguistic markers	sentiment scores of a transcript: positive value indicates positive sentiment and vice versa
ttr	788	0.66	0.16	0.21	1.00	0.5	Linguistic markers	Type-token ratio: A higher TTR means the text uses a wider variety of words relative to its length (lexically rich)
word_count_title	788	9.76	6.96	1.00	85.00	3.37	Linguistic markers	word count of a video description
word_count_transcript	788	112.99	83.33	3.00	656.00	1.28	Linguistic markers	word count of a transcription
filter	788	1.06	0.24	1.00	2.00	3.62	Platform features	binary: whether the video uses filter or not (No=1, Yes=2)
stitch_duet	788	1.05	0.22	1.00	2.00	4.15	Platform features	binary: whether the video uses stitch/duet editing features from TikTok or not (No=1, Yes=2)

4.2 Inferential statistics

Table 7 Results of GAMM on Likes, Comments and Shares

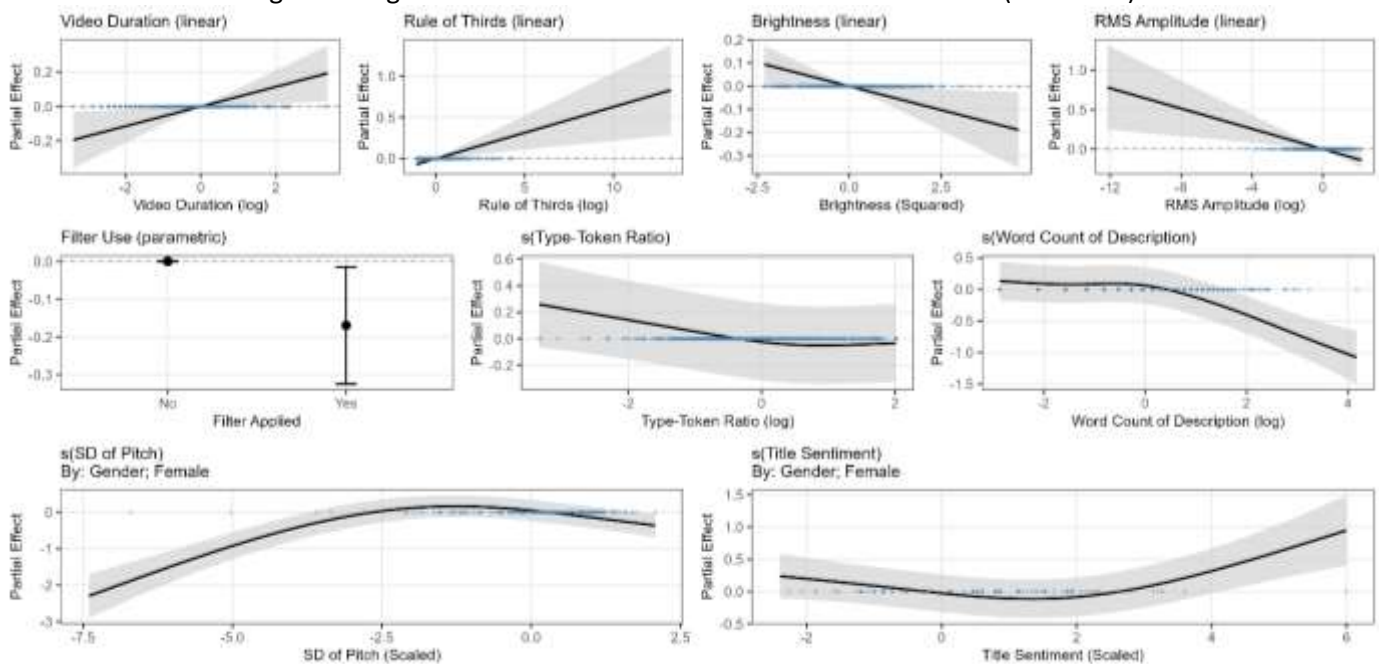
Variable/(Intercept)	Likes		Model 1b (Best)		Comments		Model 2b (Best)		Shares		Model 3b (Best)	
	Model 1a (Full)				Model 2a (Full)				Model 3a (Full)			
Parametric coefficients:												
	Estimate Significance		Estimate Significance		Estimate Significance		Estimate Significance		Estimate Significance		Estimate Significance	
(Intercept)	-2.09017	6.10E-05***	-2.4554	7.37E-11***	-5.39132	1.31E-06***	-4.78270	1.59E-11***	-7.04297	< 2e-16***	-7.44589	< 2e-16***
Views	1.00502	< 2e-16***	1.0085	< 2e-16***	-0.09584	0.0839.			-0.11240	0.1454		
Likes					1.09948	< 2e-16***	1.01370	< 2e-16***	1.29648	< 2e-16***	1.19802	< 2e-16***
Nationality: British	-0.60821	0.3332			1.90755	0.1831			1.28041	0.0467*	1.26720	0.0439*
Nationality: Chinese-Irish	-0.14313	0.9065			-0.10755	0.9691			-0.09263	0.9517	-0.06154	0.9596
filter: Yes	-0.18300	0.0322*	-0.1826	0.0206*	-0.25010	0.0341*	-0.26227	0.0167*	-0.11543	0.5195		
BGM: Yes	-0.02691	0.8615			0.04375	0.8315			-0.10358	0.9422		
sound effect: Yes	0.01280	0.9321			-0.09990	0.6251			-0.15893	0.5876		
stitch or duet: Yes	0.13173	0.1411			-0.17497	0.1529			-0.08362	0.6290		
close ties: Yes	0.01695	0.7049			0.04129	0.4982			0.32746	7.69E-06***	0.31170	2.95E-05***
Approximate significance of smooth terms:												
	edf Significance		edf Significance		edf Significance		edf Significance		edf Significance		edf Significance	
duration	1.000	0.03317*	1.000	0.0519000.	2.598	0.035881*	2.616	1.25E-02*	1.000	0.09768.	1.000	0.039324*
rule of thirds alignment	1.000	0.00334**	1.005	0.0018800**	2.917	0.181166			1.000	0.38468		
brightness	1.000	0.01744*	1.000	0.0265100*	1.000	0.338969			1.000	0.98238		
contrast	1.000	0.57260			1.326	0.823102			1.176	0.06630.	1.876	0.015872*
RMS amplitude	1.000	0.01291*	1.000	0.0077000**	1.000	0.795227			1.000	0.92861		
speech rate (syllables/s): female	1.000	0.57172			1.564	0.171036			1.000	0.02567*		
speech rate (syllables/s): male	1.000	0.86320			1.000	0.205444			1.000	0.04244*		
speech rate (syllables/s)							1.000	0.03421*			1.000	0.008642**
mean pitch: female	1.000	0.22357			1.384	0.526112			1.000	0.72094		
mean pitch: male	2.075	0.23473			1.003	0.215472			1.000	0.07991.		
standard deviation pitch: female	3.004	0.00002***	3.061	0.0000020***	1.000	0.881117	1.000	0.984778	1.000	0.67780		
standard deviation pitch: male	1.028	0.47133	1.302	0.5888000	1.000	0.001283**	1.000	0.000989***	1.000	0.03914*		
type-token ratio: female	1.000	0.02009*			2.647	0.002190**	2.636	0.003013**	1.000	0.75347	1.000	0.911597
type-token ratio: male	2.473	0.03039*			1.000	0.952701	1.000	0.677979	1.000	0.00944**	1.000	0.000399***
type-token ratio			2.093	0.0340000*								
average lexical frequency: female	1.000	0.93926			1.001	0.198091			1.552	0.71022		
average lexical frequency: male	1.000	0.30952			1.727	0.389201			1.235	0.11552		
video description's length	3.346	< 2e-16***	3.477	< 2e-16***	1.000	0.000798***	1.000	0.000282***	2.082	0.00037***	2.453	0.000004***
transcript sentiment: female	1.720	0.09873.			3.011	0.001828**	2.983	0.000664***	1.000	0.19057		
transcript sentiment: male	1.000	0.31224			1.770	0.466406	1.899	0.373690	1.000	0.17689		
transcript sentiment												
video description's sentiment: female	2.696	0.03293	2.583	0.0325100*	2.849	0.133834			1.765	0.60873		
video description's sentiment: male	1.000	0.53993	1.000	0.5183500	1.000	0.582126			3.168	0.10719		
video description's sentiment							3.178	0.055495.			1.000	0.040089*
first-person pronouns' ratio	1.242	0.75703			1.000	0.205004			1.000	0.17191		
second-person pronouns' ratio	1.000	0.25995			1.619	0.130096			1.404	0.82524		
zero crossing rate)	1.000	0.34114			1.000	0.111929			2.499	0.12636		
Adjusted R-square	0.750		0.759		0.433		0.469		0.629		0.645	
AIC	1198.400		1192.462		1771.237		1656.188		2079.934		1951.272	
N	741		741		741		741		741		741	

Note: Parametric terms, including categorical predictors and control variables, are reported with their regression coefficients (β) and p-values. Smooth terms are reported with their estimated degrees of freedom (edf) and approximate significance values, as these terms do not yield standard regression coefficients regardless of whether their estimated relationship is linear (edf = 1) or curvilinear (edf > 1).

4.2.1 Like Count

The comprehensive model (Model 1a: Likes Full Model, Table 3) incorporated 16 smooth terms and achieved an adjusted R^2 of 0.750 (AIC = 1198.40). Diagnostic checks identified numerous non-significant terms and potential basis dimension constraints, particularly for title word count (k -index = 0.63, $p < .001$) and several sentiment and pitch measures. Following the stepwise removal procedure described in Section 3.3, the parsimonious model (Model 1b: Likes Best Model, Table 3) retained eleven smooth terms, including two gender-differentiated effects and two categorical variables. This model accounted for approximately 76% of variance in log-transformed like counts (adjusted $R^2 = 0.759$, AIC = 1192.46). Smooth effects for all retained predictors are displayed in Figure 4.

Figure 18 Significant Smooth and Parametric Terms of Model 1b (Likes Best)



Note. Smooth terms with edf = 1 are effectively linear; they appear as straight lines in the figure.
SD of Pitch (Scaled): standard deviation of fundamental frequency (pitch) after scaling transformation.
Type-Token Ratio (log): log-transformed type-token ratio, indicating lexical diversity.
Word Count of Description (log): log-transformed word count of the video description.
Transcript Sentiment (Scaled): scaled sentiment score derived from the spoken transcript.
Title Sentiment (Scaled): scaled sentiment score of the video title.
Speech Rate in Syllables (log): log-transformed speech rate (syllable/s).

View counts emerged as the strongest predictor ($\beta = 1.009$, $p < .001$), with the near-unity coefficient indicating an essentially proportional relationship on the logarithmic scale. Visual composition features also proved consequential. Rule-of-thirds alignment demonstrated a positive linear effect (edf = 1.00, $p = .002$), indicating that compositionally balanced framing is associated with higher like counts. Brightness showed a significant

negative linear effect ($\text{edf} = 1.00$, $p = .027$). This variable was entered after quadratic transformation and standardisation to address skewness; the negative coefficient indicates that higher transformed brightness values were associated with fewer likes. Video duration displayed a marginally significant positive relationship with like counts ($\text{edf} = 1.00$, $p = .052$).

Audio properties also reached significance in the parsimonious model. RMS amplitude showed a significant negative effect ($\text{edf} = 1.00$, $p = .008$), with higher audio loudness associated with fewer likes. A gender-differentiated effect was observed for pitch standard deviation. For female creators specifically, pitch standard deviation exhibited strong non-linear effects ($\text{edf} = 3.06$, $p < .001$). Higher pitch variation was initially associated with more likes, with the effect levelling at higher values (see Figure 4).

Amongst linguistic and platform features, video description length exhibited the most substantial effect ($\text{edf} = 3.48$, $p < .001$), with pronounced non-linearity across its range. Type-token ratio showed a significant overall non-linear effect ($\text{edf} = 2.09$, $p = .034$). Filter use was negatively associated with like counts ($\beta = -0.183$, $p = .021$). Video description sentiment also showed a significant gender-differentiated effect, with a curvilinear pattern observed specifically for female creators ($\text{edf} = 2.58$, $p = .033$).

4.2.2 Comment count

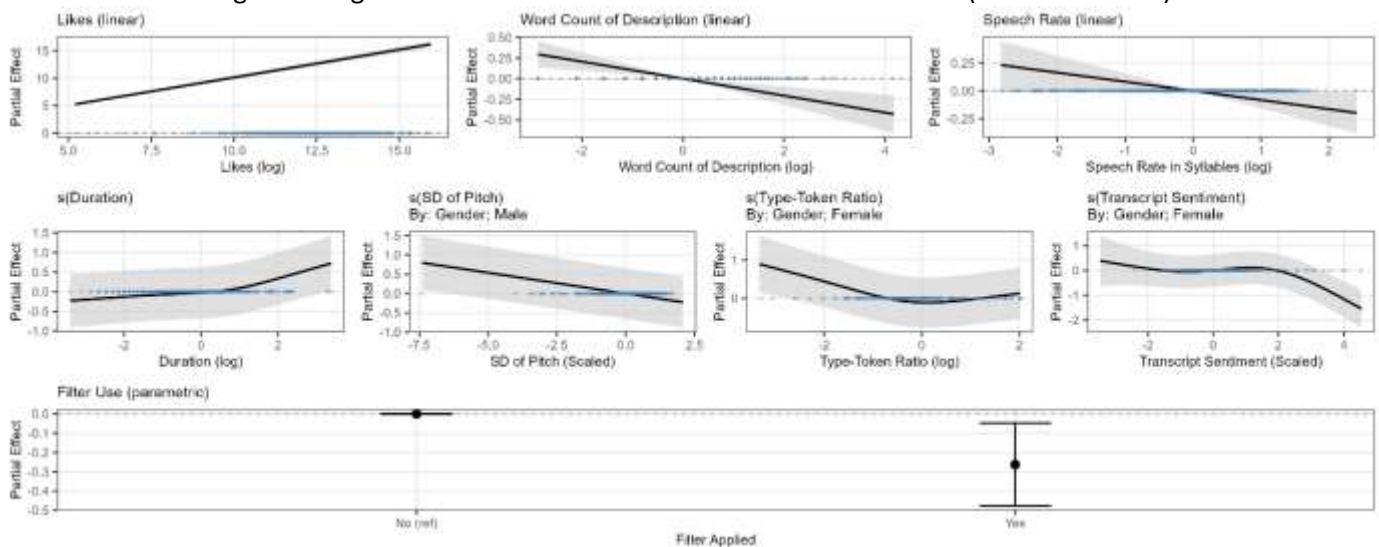
The full model (Model 2a: Comments Full Model, Table 3) included both views and likes as parametric controls to assess redundancy. Views showed only marginal significance ($\beta = -0.096$, $p = .084$), whilst likes proved substantially stronger ($\beta = 1.099$, $p < .001$). Likes subsumed the explanatory contribution of views and was accordingly retained as the sole control. The parsimonious model (Model 2b: Comments Best Model, Table 3) achieved an adjusted R^2 of 0.469 ($\text{AIC} = 1656.19$), representing an AIC reduction of 115 units from the comprehensive model ($\text{AIC} = 1771.24$). Effects for all retained predictors are displayed in Figure 5.

Likes exerted a strong positive effect on comment counts ($\beta = 1.014$, $p < .001$). Each proportional increase in likes corresponded to a roughly equivalent increase in comments. Video duration exhibited significant non-linear effects ($\text{edf} = 2.62$, $p = .012$), with comments initially increasing with duration before plateauing. Speech rate showed a significant negative effect ($\text{edf} = 1.00$, $p = .034$), indicating that faster-paced delivery is associated with fewer comments. Word count of description showed a significant negative effect ($\text{edf} = 1.00$, p

< .001), with longer descriptions associated with fewer comments even after controlling for likes. Description sentiment displayed marginally significant non-linear effects ($\text{edf} = 3.18$, $p = .056$).

Filter use was negatively associated with comment counts ($\beta = -0.262$, $p = .017$), consistent with the pattern observed for likes. The creator-level random intercept accounted for substantial variance (scale estimate = 0.419), indicating persistent individual differences in commenting behaviour beyond measured content features. Several effects were further moderated by creator gender. For male creators, pitch standard deviation showed a significant effect ($\text{edf} = 1.00$, $p < .001$). Female creators exhibited significant effects for type-token ratio ($\text{edf} = 2.64$, $p = .003$) and transcript sentiment ($\text{edf} = 2.98$, $p < .001$).

Figure 19 Significant Smooth and Parametric Terms of Model 2b (Comments Best)



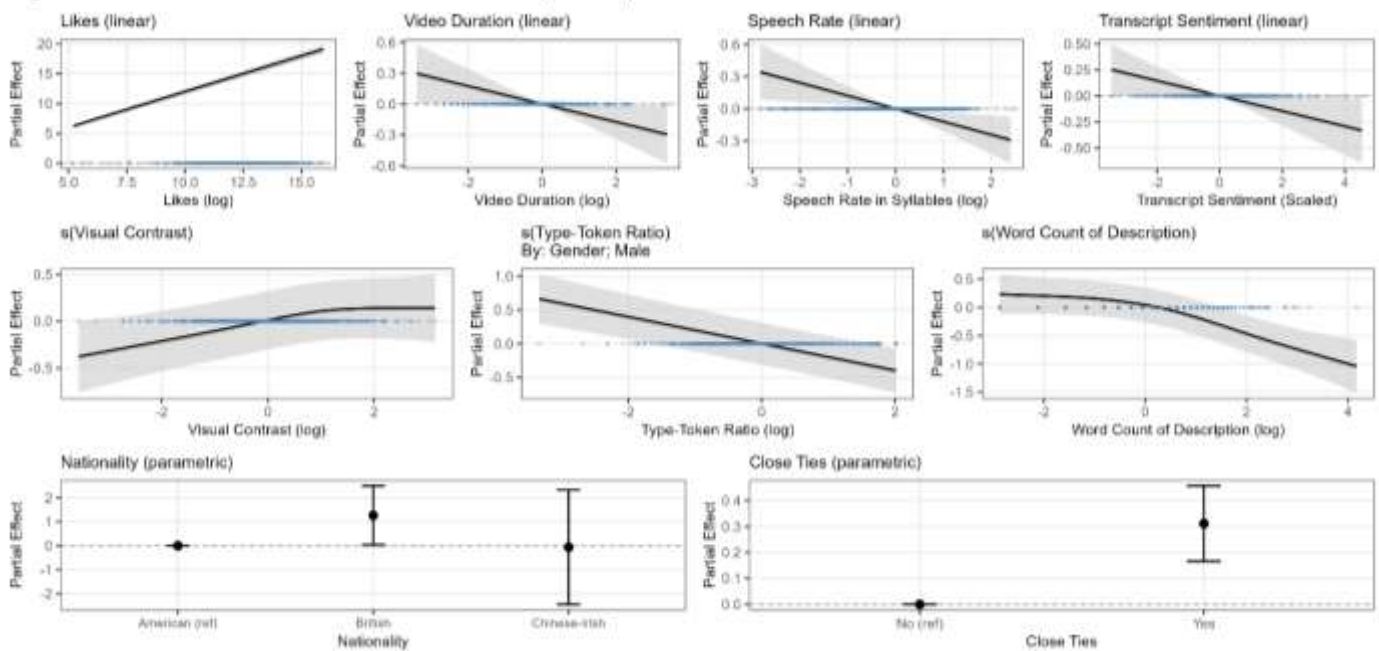
Note. Smooth terms with $\text{edf} = 1$ are effectively linear; they appear as straight lines in the figure.
SD of Pitch (Scaled): standard deviation of fundamental frequency (pitch) after scaling transformation.
Type-Token Ratio (log): log-transformed type-token ratio, indicating lexical diversity.
Word Count of Description (log): log-transformed word count of the video description.
Transcript Sentiment (Scaled): scaled sentiment score derived from the spoken transcript.
Title Sentiment (Scaled): scaled sentiment score of the video title.
Speech Rate in Syllables (log): log-transformed speech rate (syllable/s).

4.2.3 Share count

The comprehensive model (Model 3a: Shares Full Model, Table 3) achieved overall fit of adjusted $R^2 = 0.629$ ($\text{AIC} = 2079.93$). Views showed non-significant effects ($\beta = -0.112$, $p = .145$), whilst likes proved substantially stronger ($\beta = 1.296$, $p < .001$). High collinearity between views ($r = .72$) and likes ($r = .76$) with shares confirmed the decision to retain likes as the primary control. The parsimonious model (Model 3b: Shares Best Model, Table 3) retained six smooth terms, including one gender interaction, alongside three parametric

predictors. This specification achieved comparable fit (adjusted $R^2 = 0.645$, $AIC = 1951.27$) whilst substantially reducing model complexity. Effects for all retained predictors are displayed in Figure 6.

Figure 20 Significant Smooth and Parametric Tems of Model 3b (Shares Best)



Note. Smooth terms with edf = 1 are effectively linear; they appear as straight lines in the figure.
SD of Pitch (Scaled): standard deviation of fundamental frequency (pitch) after scaling transformation.
Type-Token Ratio (log): log-transformed type-token ratio, indicating lexical diversity.
Word Count of Description (log): log-transformed word count of the video description.
Transcript Sentiment (Scaled): scaled sentiment score derived from the spoken transcript.
Title Sentiment (Scaled): scaled sentiment score of the video title.
Speech Rate in Syllables (log): log-transformed speech rate (syllable/s).

Likes exerted a strong positive effect on share counts ($\beta = 1.198$, $p < .001$). Video duration showed a significant negative effect ($\text{edf} = 1.00$, $p = .039$), indicating that longer videos are associated with fewer shares. Visual contrast demonstrated a curvilinear association with shares ($\text{edf} = 1.88$, $p = .016$), with shares initially increasing with contrast level before levelling at higher values. Speech rate showed a significant negative effect ($\text{edf} = 1.00$, $p = .009$), indicating that faster-paced delivery is associated with fewer shares. Word count of description exhibited strong curvilinear effects ($\text{edf} = 2.45$, $p < .001$), with shares increasing initially before plateauing at extreme lengths. Transcript sentiment showed a significant negative effect ($\text{edf} = 1.00$, $p = .040$), indicating that content with lower sentiment scores is associated with more sharing.

Close relationship content was significantly associated with higher share counts ($\beta = 0.312$, $p < .001$). British nationality was associated with marginally higher share rates ($\beta = 1.267$, $p = .044$). The creator-level random intercept accounted for substantial variance (scale estimate = 0.676), indicating persistent individual differences in shareability beyond measured content characteristics.

Type-token ratio was the only gender-moderated effect in the shares model. It showed a significant effect for male creators specifically ($\text{edf} = 1.00$, $p < .001$). Lexical diversity thus influences sharing behaviour for audiences of male creators but not female creators.

Discussion

This study examined how multimodal content features jointly predict likes, comments, and shares in TikTok comedy videos during the COVID-19 pandemic, and how creator gender moderates these effects. Table 5 summarises the significant predictors and their directional effects across all three engagement metrics.

Three interpretive themes guide the discussion. The first concerns why likes, comments, and shares reflect qualitatively distinct audience processes rather than equivalent engagement proxies. The second examines how creator gender systematically moderates multimodal effects. The third considers what these findings contribute to genre-sensitive multimodal theory and practice.

5.1 Three Forms of Engagement as Distinct Psychological Processes

The evidence is clear but not simple. Multimodal features do jointly predict engagement, but their relative weight, functional form, and direction vary substantially across metrics and by creator gender. Like counts showed the highest explained variance, driven most strongly by visual composition, audio properties, and gender-specific prosodic features. Comment counts were most reliably shaped by temporal structure, speech pace, and the linguistic clarity of descriptions. Share counts were most distinctively predicted by relational content, video brevity, and vocal delivery.

Clearly, across all three models, video description length emerged as the single most consistent predictor, though its form and direction differ by metric. For likes and shares, it follows an inverted-U pattern: moderate description length maximises engagement, with both very short and very long descriptions reducing it. For comments, the effect is linearly negative: longer descriptions are associated with fewer comments. Our results go beyond earlier findings (Chen et al., 2021) by showing that this metric-specific pattern is non-linear and robust after accounting for a range of related factors.

Prior work, including Ling et al.'s (2021) machine learning analysis, was necessarily restricted to likes because shares and comments were unavailable in that dataset. The present study, with access to all three metrics, reveals that each reflects a qualitatively distinct audience process. Likes function as a sensory verdict: an immediate response to what is seen and heard. Comments function as a cognitive verdict: a sign that the viewer has processed the content and found it worth engaging with. Shares function as a social endorsement: a redistribution decision that reflects perceived interpersonal value. What drives each differs systematically, as Table 4 confirms.

Table 8 Summary of Significant Multimodal Predictors Across Engagement Metrics

Predictor	Likes	Comments	Shares
<i>Control</i>			
Views	↑ ***		
Likes		↑ ***	↑ ***
<i>Visual</i>			
Rule-of-thirds alignment	↑ **		
Brightness	↓ *		
Visual contrast			— *
<i>Audio</i>			
RMS amplitude	↓ **		
Speech rate		↓ *	↓ **
Video duration	↑ (marginal)	— *	↓ *
<i>Linguistic</i>			
Description length	— ***	↓ ***	— ***
Type-token ratio	— *		
Title sentiment		↓ (marginal)	
Transcript sentiment			↓ *
Close ties content			↑ ***
<i>Demographic</i>			
British nationality			↑ *
<i>Platform features</i>			
Filter use	↓ *	↓ *	
<i>Gender-moderated effects</i>			
SD of pitch (female)	— ***		
Description sentiment (female)	— *		
SD of pitch (male)		↓ ***	
Type-token ratio (female)		— **	
Transcript sentiment (female)		— ***	
Type-token ratio (male)			↓ ***

Note. ↑ = positive effect; ↓ = negative effect; — = non-linear or curvilinear effect; * = $p < .05$; ** = $p < .01$; *** = $p < .001$; marginal = $p < .10$. For likes and shares, description length follows an inverted-U pattern; for comments, the effect is linearly negative. Gender-moderated effects are estimated separately for female and male creators.

Likes can be understood as a sensory verdict, formed before much deliberation occurs. View counts emerge as the strongest predictor, replicating Ling et al.'s (2021) finding that audience reach provides the foundational baseline for passive approval. Visual composition proves distinctively consequential. Rule-of-thirds alignment produces a positive effect, and higher brightness values are associated with fewer likes, indicating that visual

aesthetics shape first impressions. In contrast with commercial contexts where louder audio benefits engagement (Xiao et al., 2025), higher RMS amplitude reduces likes in comedy. Sonic restraint appears to support comedic timing rather than attract attention. Filter use consistently reduces likes, providing GAMM-based evidence that the authenticity penalty associated with visual enhancement is robust after controlling for creator-level heterogeneity (Agrawal, 2024; Atayeva, 2024).

Two effects were gender-specific. Pitch standard deviation showed strong non-linear effects for female creators, with higher variation initially associated with more likes before levelling at extreme values. Video description sentiment also demonstrated a curvilinear pattern specifically for female creators.

Comments tell a different story. Only filter use, description length, and video duration are shared predictors with likes. The divergence reflects a fundamental difference in what commenting requires. Liking is primarily sensory: a viewer can respond positively to a well-composed, sonically restrained video without fully processing the dialogue. Commenting requires comprehension. The viewer must have engaged with the content sufficiently to formulate a response (Zhao et al., 2026).

Duration exhibits a non-linear positive effect on comments, increasing up to a plateau. Videos require sufficient substance to provoke commentary, but additional length beyond that point yields diminishing returns. Faster-paced delivery is associated with fewer comments. Rapid delivery may compress the cognitive space needed for response formation, leaving viewers with less interpretive room. Longer descriptions are also associated with fewer comments, potentially because detailed descriptions pre-empt the clarification-seeking that drives much commenting. These two patterns point in the same direction: commenting is suppressed when content is either too fast to follow or too fully explained to invite elaboration.

Shared emotion is another pathway to comment engagement. Zhao et al. (2026) identify empathy evocation as the primary mechanism by which emotional expression generates attention on social media. Diverse emotional content resonates with a broader audience and elicits more responses. Consistent with this, description sentiment showed marginal non-linear effects on comments. Transcript sentiment showed a significant effect specifically for female creators, suggesting that the affective tone of spoken content shapes commenting behaviour in a gender-differentiated way.

For gender-specific effects on comments, pitch standard deviation showed a significant positive effect for male creators. Female creators also exhibited significant effects for type-token ratio.

Sharing is a social endorsement, and the predictor profile for shares reflects this most clearly. The overlap with likes and comments is limited. Redistribution decisions appear governed by perceived interpersonal value rather than sensory appreciation or cognitive processing alone.

The clearest evidence comes from close ties content. Videos focused on family, friendship, and personal connection are significantly more likely to be shared, even after controlling for prior popularity (Chung et al., 2023). Content perceived as interpersonally meaningful crosses a threshold that merely entertaining content may not. Several other patterns are consistent with this logic. Shorter videos attract more shares: brevity likely reduces the effort threshold for forwarding. Slower delivery is also associated with more shares, perhaps because more deliberate pacing makes content feel recommendable. Transcript sentiment shows a negative effect: content with lower sentiment scores accrues more shares. In comedy, sardonic or emotionally ambivalent material may carry greater social currency than uniformly positive content. This contrasts with Berger and Milkman's (2012) finding that emotionally arousing content promotes sharing, suggesting that sentiment functions differently in comedy than in other contexts. Description length shows the same inverted-U pattern as for likes. Visual contrast shows a curvilinear positive association, initially increasing shareability before levelling. These non-linearities would have been undetectable under a linear or prespecified quadratic approach (Xiao et al., 2025). British nationality was also associated with marginally higher share rates.

Type-token ratio was the only gender-moderated effect in the shares model. It showed a significant effect for male creators specifically, with higher lexical diversity associated with fewer shares.

Across all three models, the predictor profiles are sufficiently distinct to warrant treating likes, comments, and shares as separate dependent variables. Future research should avoid collapsing them into a composite engagement score or modelling only one as a virality proxy.

5.2 Gender as a Systematic Moderator

Gender moderated multimodal effects across all three engagement models. Our findings show that male and female creators deploy distinct multimodal strategies in audience-directed communication, extending work on

gendered broadcasting styles (cf. Zhang et al., 2025) to the comedy context. For female creators, prosodic expressiveness drove likes, whilst linguistic complexity and emotional tone drove comments. For male creators, pitch variability influenced comments, whilst lexical properties drove shares.

Our study has important implications for multimodal and engagement research, particularly for studies that treat content features as uniform in effect. Unlike Xiao et al.'s (2025) short sales video study, which did not incorporate gender as a moderating variable, the present findings reveal that the same acoustic feature, pitch standard deviation, predicts likes for female creators but comments for male creators. A gender-pooled model would have obscured this distinction entirely. Future multimodal engagement research should treat gender as a default moderator rather than an optional variable.

The gender interactions further suggest that audience responses are shaped not only by what creators do, but by the gendered communicative expectations through which those behaviours are interpreted. Sociocultural norms establish different expectations for vocal behaviour by gender (Leyns et al., 2024; Zhang et al., 2025), and audiences appear to respond accordingly. Whether these patterns generalise across cultural contexts with different gender norms remains an open question. Still, the consistency of the moderation pattern across all three metrics suggests that gendered interpretation is a robust feature of audience response in short-form comedy.

Creator demographics demonstrated limited direct predictive power across engagement metrics. Nationality reached significance only in the shares model, where British creators showed marginally higher share rates. Yet the necessity of creator-specific random intercepts in all three models indicates substantial between-creator variance in baseline engagement. This finding is in line with Ling et al.'s (2021) finding that creator popularity is the single most powerful predictor of virality, and extends Bhandari and Bimo's (2022) observation that creator attributes shape content style and audience reception to show that multimodal effects operate within a context already structured by social capital.

5.3 Genre Specificity and Comparison with Prior Multimodal Research

A theoretically important feature of the present findings is the degree to which they diverge from prior multimodal engagement research conducted in commercial or educational content contexts. These divergences reveal how the same multimodal cues acquire different predictive valence depending on genre.

Background music and sound effects showed no significant effects across any engagement metric. This contrasts with Wahid et al.'s (2022) finding that original audio enhances engagement across content genres. In comedy, sonic minimalism appears to serve the genre: dialogue is foregrounded rather than competing with musical accompaniment. A direct comparison with Xiao et al.'s (2025) short sales video study is particularly instructive. That study found significant effects of audio spectral features, including fluctuant frequency and pitch, uniformly across all three engagement outcomes. The present study also finds pitch effects for comedy, but these are gender-moderated and metric-specific rather than uniform. In sales contexts, audio signals carry commercial information and signal product quality. In comedy, they carry comedic intent and emotional investment. The same acoustic property serves fundamentally different communicative functions depending on genre.

The GAMM framework also offers analytical advantages over the approach used by Xiao et al. (2025), who restricted non-linearity analysis to prespecified quadratic terms and did not account for non-independence of videos from the same creator. The present study addresses both limitations through flexible spline smoothing and creator-level random intercepts. The inverted-U effects of description length on likes and shares illustrate the analytical payoff: this pattern would have been obscured under a linear or prespecified quadratic approach. Prior multimodal studies typically employ regression approaches that either impose linearity or restrict detectable functional forms (Chen et al., 2024; Chung et al., 2023; Xiao et al., 2025). The GAMM approach requires neither.

Sentiment effects further illuminate genre-platform contingency. Aiger et al. (2026) report that emotional polarity accounted for only a small proportion of variance in science communication and did not significantly predict engagement overall. The present findings differ substantially. Across all three models, sentiment at the transcript or description level was a meaningful predictor, particularly for female creators and sharing behaviour. Comedy operates by subverting emotional expectation, so affective signals are constitutive of the

genre rather than incidental to it. The same cue carries different predictive weight in comedy versus science communication (Aiger et al., 2026), arguing in favour of genre-sensitive analytical frameworks rather than platform-level generalisations. Pronoun usage patterns point in the same direction: second-person pronouns significantly predicted engagement in news and informational content (Cheng & Li, 2024; Chung et al., 2023) but proved non-significant for comedy, suggesting that humour relies more on narrative distance than direct interpersonal address.

Engagement metrics carry interpretive limitations worth acknowledging. Barrie et al. (2025) caution that engagement-based measures may overstate echo chamber effects and underestimate incidental exposure. Montag et al. (2021) find that passive consumption dominates platform behaviour, and Gu et al. (2022) identify user segments characterised by minimal active engagement despite high platform use. The present findings likely capture selective amplification patterns rather than the full spectrum of audience attention dynamics. Future research integrating watch time, replays, and scrolling pauses would provide more comprehensive accounts.

5.4 Theoretical, Methodological, and Practical Implications

This study makes three theoretical contributions. First, the findings advance multimodal communication theory by demonstrating that engagement emerges through complex and asymmetric interactions between visual, audio, linguistic, and platform-specific features, rather than any single modality in isolation. This addresses the fragmentation critique levelled at prior research (Chen, 2024; Wang, 2021). Second, the findings support a differentiated model of engagement wherein distinct metrics capture distinct psychological processes. This aligns with calls to move beyond engagement-as-metric towards engagement-as-process (Gu et al., 2022; Montag et al., 2021) and provides empirical grounding for that theoretical shift. Third, the GAMM framework with creator-specific random intercepts represents a methodological contribution. It simultaneously accommodates non-linear relationships, hierarchical data structures, and creator-level heterogeneity. Prior multimodal studies either impose linearity or restrict detectable functional forms (Chen et al., 2024; Chung et al., 2023; Xiao et al., 2025). The GAMM approach requires neither. The observational design, whilst precluding

definitive causal inference, captures engagement as it genuinely occurs and offers ecological validity that laboratory experiments cannot replicate.

For content creators, the findings point toward a metric-specific content strategy. A creator seeking likes will optimise visual composition and restrain audio intensity. One seeking comments will invest in substantive, moderately paced videos with emotionally legible content, keeping descriptions concise to preserve discussion space. One seeking shares will prioritise relational themes and edgier affective content, keeping videos brief and pacing delivery deliberately. None of these prescriptions applies universally. Yet the underlying principle holds: treating all engagement metrics as equivalent targets produces content optimised for none of them. The differentiated predictive profiles documented here offer a more principled basis for content decisions than aggregate engagement scores. Female creators, in particular, may benefit from attending to prosodic expressiveness as a driver of passive approval. Male creators, by contrast, may find that attending to lexical simplicity is particularly consequential for share amplification, more so than for comments or likes.

Conclusion

This study examined how visual, acoustic, linguistic, and platform-specific features jointly predict likes, comments, and shares in TikTok comedy videos. Three principal findings emerge. Engagement metrics are not interchangeable. Likes reflects low-effort sensory responses shaped by visual composition, audio restraint, and description length. Comments reflect interpretive participation driven by duration, linguistic complexity, and speech pace. Shares reflect social endorsement shaped by relational content, sentiment, and gender-specific lexical properties.

Genre specificity matters, and its effects run deeper than expected. Filter use, sonic minimalism, and sentiment operate in comedy-specific ways that diverge markedly from findings in commercial and educational content contexts. These divergences reinforce the necessity of domain-specific investigations rather than assuming uniform multimodal effects across content categories. Creator identity operates through accumulated social capital rather than demographic attributes alone, whilst gender consistently moderates multimodal cue effects across all three metrics.

Future work should integrate attention-based measures such as watch time and replays. Experimental designs would strengthen causal interpretation, and extending the analytical framework to other genres and platforms would test the boundary conditions of genre-specific effects. Short-form video platforms continue to shape public discourse well beyond entertainment. Ongoing societal concerns include algorithmic bias and representational harm (Ionescu & Licu, 2023), and engagement mechanisms have been linked to misinformation amplification and the reinforcement of echo chambers (González-Esteban et al., 2024; Rejeb et al., 2023; Solangi, 2023). How multimodal features interact with these algorithmic and societal forces remains underexplored. The present study provides a foundation for that work.

References (Chapter 2):

- Abdel-Hamid, L., Shaker, N. H., & Emara, I. (2020). Analysis of linguistic and prosodic features of bilingual arabic–english speakers for speech emotion recognition. *IEEE Access*, 8, 72957–72970.
<https://doi.org/10.1109/ACCESS.2020.2987864>
- Agabo, S. (2024). Digital activism through humour and satire: An analysis of the use of TikTok and X for political accountability in Uganda. *Theses & Dissertations*.
https://ecommons.aku.edu/theses_dissertations/2257
- Agrawal, E. (2024). *Going viral: An analysis of advertising of technology products on TikTok*. arXiv.
<https://doi.org/10.48550/ARXIV.2402.00010>
- Ahlse, J., Nilsson, F., & Sandström, N. (2020). *It's time to TikTok: Exploring Generation Z's motivations to participate in #Challenges*. <https://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-48708>
- Aiger, M., Elboj, C., Lozano-Blasco, R., & Acero-Ferrero, M. (2026). Science communication in social Media: Analysis of success on TikTok, Instagram, and YouTube across scientific disciplines. *Computers in Human Behavior*, 177, 108866. <https://doi.org/10.1016/j.chb.2025.108866>
- Al-Rawi, A., Grepin, K., Li, X., Morgan, R., Wenham, C., & Smith, J. (2021). Investigating public discourses around gender and covid-19: A social media analysis of twitter data. *Journal of Healthcare Informatics Research*, 5(3), 249–269. <https://doi.org/10.1007/s41666-021-00102-x>
- Anderson, K. E. (2020). Getting acquainted with social networks and apps: It is time to talk about TikTok. *Library Hi Tech News*, 37(4), 7–12. <https://doi.org/10.1108/LHTN-01-2020-0001>
- Angelopoulou, G., Meier, E. L., Kasselimis, D., Pan, Y., Tsolakopoulos, D., Velonakis, G., Karavasilis, E., Kelekis, N. L., Goutsos, D., Potagas, C., & Kiran, S. (2020). Investigating gray and white matter structural substrates of sex differences in the narrative abilities of healthy adults. *Frontiers in Neuroscience*, 13, 1424. <https://doi.org/10.3389/fnins.2019.01424>
- Atayeva, N. 2024. Understanding Concept of Virality in Social Media from Users' Perspective. Available from: <https://urn.fi/URN:NBN:fi:jyu-202406134593>.

Baghdadi, J. D., Coffey, K. C., Belcher, R., Frisbie, J., Hassan, N., Sim, D., & Malik, R. D. (2023).

#Coronavirus on TikTok: User engagement with misinformation as a potential threat to public health behavior. *JAMIA Open*, 6(1), ooad013. <https://doi.org/10.1093/jamiaopen/ooad013>

Bai, D., & Gu, Y. (2026). Harnessing Big Data, Hindered by Bias: Evaluating TikTok Research API for Fair and Optimal Social Sciences. *Social Science Computer Review*, 08944393251413277. <https://doi.org/10.1177/08944393251413277>

Balingbing, C., Kirchner, S., Siebald, H., Van Hung, N., & Hensel, O. (2024). Determining the sound signatures of insect pests in stored rice grain using an inexpensive acoustic system. *Food Security*, 16(6), 1529–1538. <https://doi.org/10.1007/s12571-024-01493-6>

Barrie, C., Atalay, A., & ElKattan, A. (2025). *Information diets are more diverse in attention than in engagement*. https://doi.org/10.31219/osf.io/kvbzd_v3

Barta, S., Belanche, D., Fernández, A., & Flavián, M. (2023). Influencer marketing on TikTok: The effectiveness of humor and followers' hedonic experience. *Journal of Retailing and Consumer Services*, 70, 103149. <https://doi.org/10.1016/j.jretconser.2022.103149>

Barthelme, S. (2015). *Imager: Image processing library based on 'cimg'* [Data set]. <https://doi.org/10.32614/CRAN.package.imager>

Basch, C. H., Fera, J., Pierce, I., & Basch, C. E. (2021). Promoting mask use on tiktok: Descriptive, cross-sectional study. *JMIR Public Health and Surveillance*, 7(2), e26392. <https://doi.org/10.2196/26392>

Bhandari, A., & Bimo, S. (2022). Why's Everyone on TikTok Now? The Algorithmized Self and the Future of Self-Making on Social Media. *Social Media + Society*, 8(1), 205630512210862. <https://doi.org/10.1177/20563051221086241>

Bilozerskyi, V., Dergachov, K., Krasnov, L., Zymovin, A., & Popov, A. (2023). New method for video stream brightness stabilization: Algorithms and performance evaluation. *Radioelectronic and Computer Systems*, (3), 125–135. <https://doi.org/10.32620/reks.2023.3.10>

Boeker, M., & Urman, A. (2022). An empirical investigation of personalization factors on TikTok. *Proceedings of the ACM Web Conference 2022*, 2298–2309. <https://doi.org/10.1145/3485447.3512102>

- Cervi, L., & Divon, T. (2023). Playful activism: Memetic performances of Palestinian resistance in Tiktok #challenges. *Social Media + Society*, 9(1), 20563051231157607. <https://doi.org/10.1177/20563051231157607>
- Chen, J. (2024). Application of audiovisual language on the TikTok platform and analysis of its communication effect. *Advances in Social Behavior Research*, 11, 33–40. <https://doi.org/10.54254/2753-7102/11/2024104>
- Chen, K., Song, N., Zhao, Y., Peng, J., & Chen, Y. (2024). Online attention versus knowledge utilization: Exploring how linguistic features of scientific papers influence knowledge diffusion. *Information Processing & Management*, 61(3), 103691. <https://doi.org/10.1016/j.ipm.2024.103691>
- Chen, Q., Min, C., Zhang, W., Ma, X., & Evans, R. (2021). Factors driving citizen engagement with government tiktok accounts during the covid-19 pandemic: Model development and analysis. *Journal of Medical Internet Research*, 23(2), e21463. <https://doi.org/10.2196/21463>
- Chen, X. (2024). Investigation on the self-improving algorithm of TikTok based on extensive user interactions: *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence*, 227–233. <https://doi.org/10.5220/0012925100004508>
- Cheng, Z., & Li, Y. (2024). Like, comment, and share on Tiktok: Exploring the effect of sentiment and second-person view on the user engagement with Tiktok news videos. *Social Science Computer Review*, 42(1), 201–223. <https://doi.org/10.1177/08944393231178603>
- Chłopicki, W., & Brzozowska, D. (2021). Sophisticated humor against COVID-19: The Polish case. *HUMOR*, 34(2), 201–227. <https://doi.org/10.1515/humor-2021-0015>
- Chung, J. (Jae), Ding, Y., & Kalra, A. (2023). I really know you: How influencers can increase audience engagement by referencing their close social ties. *Journal of Consumer Research*, 50(4), 683–703. <https://doi.org/10.1093/jcr/ucad019>
- Collins, A. (2024). Joke economics: The low profile of comedy in the economics of arts and culture. *Journal of Cultural Economics*, 48(3), 339–346. <https://doi.org/10.1007/s10824-024-09504-2>

- Damhoureyeh, M. A., Darawsheh, W. B., Qa'dan, W. N., & Natour, Y. S. (2020). Preliminary speech rate normative data in adult jordanian speakers. *Journal of Language Teaching and Research*, 11(2), 204. <https://doi.org/10.17507/jltr.1102.08>
- Demina, M. A. (2021). An investigation into gender specifics of communicants' pitch attunement in natural English conversation. *XLinguae*, 14(2), 142–156. <https://doi.org/10.18355/XL.2021.14.02.11>
- Dong, Y., Liu, S., Ma, C., & Zhou, Z. (2023). Stability analysis of loess landslide caused by cutting slope excavation based on extension theory. *11th Academic Conference of Geology Resource Management and Sustainable Development 2023*, 1104–1109. <https://doi.org/10.52202/073371-0138>
- Duffy, B. E., Poell, T., & Nieborg, D. B. (2019). Platform practices in the cultural industries: Creativity, labor, and citizenship. *Social Media + Society*, 5(4), 2056305119879672. <https://doi.org/10.1177/2056305119879672>
- Eide, A.G. and Jonassen, V.A. 2024. *Master Thesis on Virality and TikTok*. (Online) Master thesis, Handelshøyskolen BI. (Accessed 31 July 2025). Available from: <https://biopen.bi.no/bi-xmlui/handle/11250/3165789>.
- Er, M. B. (2020). A novel approach for classification of speech emotions based on deep and acoustic features. *IEEE Access*, 8, 221640–221653. <https://doi.org/10.1109/ACCESS.2020.3043201>
- Erdoğan, F., & Çakıroğlu, Ü. (2021). The educational power of humor on student engagement in online learning environments. *Research and Practice in Technology Enhanced Learning*, 16(1), 9. <https://doi.org/10.1186/s41039-021-00158-8>
- Eslami, S. P., Ghasemaghaci, M., & Hassanein, K. (2022). Understanding consumer engagement in social media: The role of product lifecycle. *Decision Support Systems*, 162, 113707. <https://doi.org/10.1016/j.dss.2021.113707>
- Feng, X., Xiong, J., & Tang, Z. (2023). To be rational or sensitive? The gender difference in how textual environment cue and personal characteristics influence the sentiment expression on social media. *Telematics and Informatics*, 80, 101971. <https://doi.org/10.1016/j.tele.2023.101971>

- Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models* (1st edn). Cambridge University Press. <https://doi.org/10.1017/CBO9780511790942>
- González-Esteban, J.-L., Lopez-Rico, C. M., Morales-Pino, L., & Sabater-Quinto, F. (2024). Intensification of hate speech, based on the conversation generated on tiktok during the escalation of the war in the middle east in 2023. *Social Sciences*, 13(1), 49. <https://doi.org/10.3390/socsci13010049>
- Grzenkowicz, M., & Wildfeuer, J. (2025). Addressing TikTok's multimodal complexity: A multi-level annotation scheme for the audio-visual design of short video content. *Digital Scholarship in the Humanities*, fqaf047. <https://doi.org/10.1093/llc/fqaf047>
- Gu, L., Gao, X., & Li, Y. (2022). What drives me to use TikTok: A latent profile analysis of users' motives. *Frontiers in Psychology*, 13, 992824. <https://doi.org/10.3389/fpsyg.2022.992824>
- Guinaudeau, B., Munger, K., & Votta, F. (2022). Fifteen Seconds of Fame: TikTok and the Supply Side of Social Video. *Computational Communication Research*, 4(2), 463–485. <https://doi.org/10.5117/CCR2022.2.004.GUIN>
- Habibi, S. A., & Salim, L. (2021). Static vs. dynamic methods of delivery for science communication: A critical analysis of user engagement with science on social media. *PLOS ONE*, 16(3), e0248507. <https://doi.org/10.1371/journal.pone.0248507>
- Heikkonen, A. 2024. Viral video marketing on TikTok: effects on brands beyond views, likes and shares. *laturi.oulu.fi*. (Online). (Accessed 31 July 2025). Available from: <https://oulurepo.oulu.fi/handle/10024/50127>.
- Hove, M. J., Vuust, P., & Stupacher, J. (2019). Increased levels of bass in popular music recordings 1955–2016 and their relation to loudness. *The Journal of the Acoustical Society of America*, 145(4), 2247–2253. <https://doi.org/10.1121/1.5097587>
- Ionescu, C. G., & Licu, M. (2023). Are tiktok algorithms influencing users' self-perceived identities and personal values? A mini review. *Social Sciences*, 12(8), 465. <https://doi.org/10.3390/socsci12080465>
- Jenkins, H., Itō, M., & boyd, danah. (2016). *Participatory culture in a networked era: A conversation on youth, learning, commerce, and politics*. Polity Press.

- Jin, X., Li, X., Lou, H., Fan, C., Deng, Q., Xiao, C., Cui, S., & Singh, A. K. (2022). Aesthetic attribute assessment of images numerically on mixed multi-attribute datasets. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 18(3s), 1–16. <https://doi.org/10.1145/3547144>
- Jockers, M. (2015). *Syuzhet: Extracts sentiment and sentiment-derived plot arcs from text* [Data set]. <https://doi.org/10.32614/CRAN.package.syuzhet>
- Katiaj, P., Farmakis, T., Kouropoulos, A. and Lounis, S. 2024. Storytelling in TikTok: A Path to Engagement. (Accessed 31 July 2025). Available from: <https://mpira.ub.uni-muenchen.de/123280/>.
- Kaye, D. B. V., Zeng, J., & Wikström, P. (2022). *TikTok: Creativity and culture in short video*. Polity.
- Klug, D. (2020). 'It took me almost 30 minutes to practice this'. *Performance and Production Practices in Dance Challenge Videos on TikTok*. <https://doi.org/10.48550/ARXIV.2008.13040>
- Klug, D., Qin, Y., Evans, M., & Kaufman, G. (2021). Trick and Please. A Mixed-Method Study On User Assumptions About the TikTok Algorithm. *13th ACM Web Science Conference 2021*, 84–92. <https://doi.org/10.1145/3447535.3462512>
- Kostyk, A., & Huhmann, B. A. (2021). Perfect social media image posts: Symmetry and contrast influence consumer response. *European Journal of Marketing*, 55(6), 1747–1779. <https://doi.org/10.1108/EJM-09-2018-0629>
- Lee, J., Im, J., Kim, D., & Nam, J. (2024). *Video-foley: Two-stage video-to-sound generation via temporal event condition for foley sound*. arXiv. <https://doi.org/10.48550/ARXIV.2408.11915>
- Leech, G. N., Rayson, P., & Wilson, A. (2014). *Word frequencies in written and spoken English: Based on the British National Corpus*. Taylor and Francis. https://ucrel.lancs.ac.uk/bncfreq/lists/2_1_spokenvwritten_alpha.txt
- Leyns, C., Papeleu, T., Feryn, T., De Baer, S., Bettens, K., Corthals, P., & D'haeseleer, E. (2024). Age and gender differences in belgian dutch intonation. *Journal of Voice*, 38(3), 801.e1-801.e26. <https://doi.org/10.1016/j.jvoice.2021.12.011>

- Li, M., Li, G., & Yang, S. (2024). Correction by distraction: How high-tempo music enhances medical experts' debunking TikTok videos. *Journal of Computer-Mediated Communication*, 29(5), zmae007. <https://doi.org/10.1093/jcmc/zmae007>
- Li, Y., Guan, M., Hammond, P., & Berrey, L. E. (2021). Communicating COVID-19 information on TikTok: A content analysis of TikTok videos from official accounts featured in the COVID-19 information hub. *Health Education Research*, 36(3), 261–271. <https://doi.org/10.1093/her/cyab010>
- Ligges, U., Krey, S., Mersmann, O., Schnackenberg, S., Guénard, G., Ellis, D. P. W., & Underbit Technologies. (2004). *Tuner: Analysis of music and speech* [Data set]. <https://doi.org/10.32614/CRAN.package.tuneR>
- Ling, C., Blackburn, J., De Cristofaro, E., & Stringhini, G. (2022). Slapping cats, bopping heads, and oreo shakes: Understanding indicators of virality in tiktok short videos. *14th ACM Web Science Conference 2022*, 164–173. <https://doi.org/10.1145/3501247.3531551>
- Lockyer, S., & De Benedictis, S. (2023). Performing pregnancy: Comic content, critique and ambivalence in pregnant stand-up comedy. *International Journal of Cultural Studies*, 26(3), 343–361. <https://doi.org/10.1177/13678779231166444>
- Lu, S., Yu, M., & Wang, H. (2023). What matters for short videos' user engagement: A multiblock model with variable screening. *Expert Systems with Applications*, 218, 119542. <https://doi.org/10.1016/j.eswa.2023.119542>
- Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, 6(2), 114–133. <https://doi.org/10.1177/1745691611400234>
- Meer, P., & Fuchs, R. (2022). The trini sing-song: Sociophonetic variation in trinidadian english prosody and differences to other varieties. *Language and Speech*, 65(4), 923–957. <https://doi.org/10.1177/0023830921998404>
- Meng, K. S., & Leung, L. (2021). Factors influencing TikTok engagement behaviors in China: An examination of gratifications sought, narcissism, and the Big Five personality traits. *Telecommunications Policy*, 45(7), 102172. <https://doi.org/10.1016/j.telpol.2021.102172>

- Ming, S., Han, J., Yao, X., Guo, X., Guo, Q., & Lei, B. (2024). Myopia information on TikTok: Analysis factors that impact video quality and audience engagement. *BMC Public Health*, 24(1), 1194. <https://doi.org/10.1186/s12889-024-18687-4>
- Mönnich, A.-L., Strieth, S., Bohnert, A., Ernst, B. P., & Rader, T. (2023). The German hearing in noise test with a female talker: Development and comparison with German male speech test. *European Archives of Oto-Rhino-Laryngology*, 280(7), 3157–3169. <https://doi.org/10.1007/s00405-023-07820-5>
- Montag, C., Yang, H., & Elhai, J. D. (2021). On the psychology of tiktok use: A first glimpse from empirical findings. *Frontiers in Public Health*, 9, 641673. <https://doi.org/10.3389/fpubh.2021.641673>
- Mullen, L. (2016). *Tokenizers: Fast, consistent tokenization of natural language text* [Data set]. <https://doi.org/10.32614/CRAN.package.tokenizers>
- Murthy, D., Ouellette, R. R., Anand, T., Radhakrishnan, S., Mohan, N. C., Lee, J., & Kong, G. (2024). Using computer vision to detect e-cigarette content in Tiktok videos. *Nicotine and Tobacco Research*, 26(Supplement_1), S36–S42. <https://doi.org/10.1093/ntr/ntad184>
- Neer, E. M., Brahmbhatt, A., Walsh, C. R., & Warlaumont, A. S. (2025). Pitch characteristics of real-world infant-directed speech vary with pragmatic context, perceived adult gender, and infant gender. *PLOS One*, 20(6), e0326569. <https://doi.org/10.1371/journal.pone.0326569>
- Ngilangil, K. M. (2022). Tiktok on SNSU students: Engagement and influence. *International Journal of English Literature and Social Sciences*, 7(4), 150–155. <https://doi.org/10.22161/ijels.74.22>
- Omar, B., & Dequan, W. (2020). Watch, Share or Create: The Influence of Personality Traits and User Motivation on TikTok Mobile Video Usage. *International Journal of Interactive Mobile Technologies (iJIM)*, 14(04), 121. <https://doi.org/10.3991/ijim.v14i04.12429>
- Ooms, J. (2016). *Magick: Advanced graphics and image-processing in r* [Data set]. <https://doi.org/10.32614/CRAN.package.magick>
- Rejeb, A., Rejeb, K., Appolloni, A., & Treiblmaier, H. (2023). Foundations and knowledge clusters in TikTok (Douyin) research: Evidence from bibliometric and topic modelling analyses. *Multimedia Tools and Applications*, 83(11), 32213–32243. <https://doi.org/10.1007/s11042-023-16768-x>

Rinker, T. (2017). *Textstem: Tools for stemming and lemmatizing text* [Data set].

<https://doi.org/10.32614/CRAN.package.textstem>

Schellewald, A. (2021). Communicative Forms on TikTok: Perspectives From Digital Ethnography.

International Journal of Communication, 15(0), 21. <https://ijoc.org/index.php/ijoc/article/view/16414>

Schwenzow, J., Hartmann, J., Schikowsky, A., & Heitmann, M. (2021). Understanding videos at scale: How to extract insights for business research. *Journal of Business Research*, 123, 367–379.

<https://doi.org/10.1016/j.jbusres.2020.09.059>

Shahbaznezhad, H., Dolan, R., & Rashidirad, M. (2021). The role of social media content format and platform in users' engagement behavior. *Journal of Interactive Marketing*, 53(1), 47–65.

<https://doi.org/10.1016/j.intmar.2020.05.001>

Shen, X., & Lu, Q. (2025). Consistency of restricted maximum likelihood estimators in high-dimensional kernel linear mixed-effects models with applications in estimating genetic heritability. *Mathematics*, 13(15), 2363. <https://doi.org/10.3390/math13152363>

Shi, A., El Haddad, J., Cai, P., Song, S., Wang, Y. J., Liu, Q., & Li, P. (2023). Mpox (Monkeypox) information on TikTok: Analysis of quality and audience engagement. *BMJ Global Health*, 8(3), e011138.

<https://doi.org/10.1136/bmjgh-2022-011138>

Shi, J., Gu, Y., & Vigliocco, G. (2023). Prosodic modulations in child-directed language and their impact on word learning. *Developmental Science*, 26(4), e13357. <https://doi.org/10.1111/desc.13357>

Shutsko, A. 2020. User-Generated Short Video Content in Social Media. A Case Study of TikTok In: G. Meiselwitz, ed. *Social Computing and Social Media. Participation, User Experience, Consumer Experience, and Applications of Social Computing* (Online). Cham: Springer International Publishing, pp.108–125. (Accessed 31 July 2025). Available from: https://link.springer.com/10.1007/978-3-030-49576-3_8.

Singh, S. (2025, May 7). How many people use TikTok 2025 (active users). *DemandSage*.

<https://www.demandsage.com/tiktok-user-statistics/>

- Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, 11(1), 54–71. <https://doi.org/10.1037/1082-989X.11.1.54>
- Solangi, S. (2023). From memes to movements: How tiktok algorithms amplify youth activism in pakistan's 2024 elections. *International Journal of Social Science Research Praxis (IJSSRP)*, 1(1), 51–62. <https://ijssrp.com/index.php/ojs/article/view/10>
- Su, L. Y.-F., McKasy, M., Cacciatore, M. A., Yeo, S. K., DeGrauw, A. R., & Zhang, J. S. (2022). Generating science buzz: An examination of multidimensional engagement with humorous scientific messages on twitter and instagram. *Science Communication*, 44(1), 30–59. <https://doi.org/10.1177/10755470211063902>
- Sueur, J., Aubin, T., & Simonis, C. (2006). *Seewave: Sound analysis and synthesis* [Data set]. <https://doi.org/10.32614/CRAN.package.seewave>
- Sun, B., Mao, H., & Yin, C. (2020). Male and female users' differences in online technology community based on text mining. *Frontiers in Psychology*, 11, 806. <https://doi.org/10.3389/fpsyg.2020.00806>
- Sun, L., Zhang, H., Zhang, S., & Luo, J. (2020). *Content-based analysis of the cultural differences between tiktok and douyin*. arXiv. <https://doi.org/10.48550/ARXIV.2011.01414>
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (Seventh edition). Pearson.
- UK annex | TikTok. (2025, July). TikTok. <https://www.tiktok.com/legal/page/global/uk-annex/en>
- Van Ee, R., Van Boxtel, J. J. A., Parker, A. L., & Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. *The Journal of Neuroscience*, 29(37), 11641–11649. <https://doi.org/10.1523/JNEUROSCI.0873-09.2009>
- Vásquez, C., & Aslan, E. (2021). “Cats be outside, how about meow”: Multimodal humor and creativity in an internet meme. *Journal of Pragmatics*, 171, 101–117. <https://doi.org/10.1016/j.pragma.2020.10.006>
- Vijay, D. and Gekker, A. 2021. Playing Politics: How Sabarimala Played Out on TikTok. *American Behavioral Scientist*. 65(5), pp.712–734.

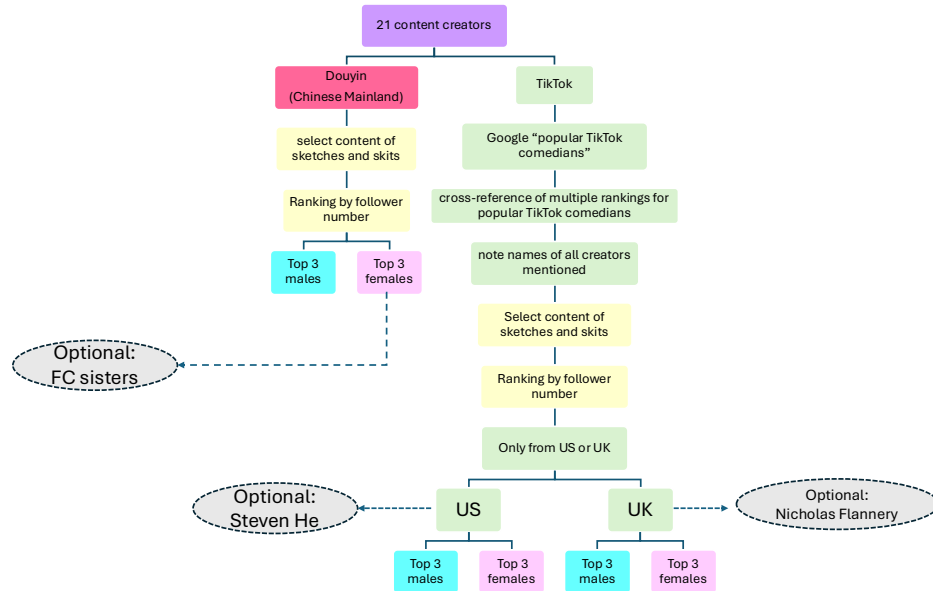
- Vinerean, S., & Opreana, A. (2021). Measuring customer engagement in social media marketing: A higher-order model. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(7), 2633–2654. <https://doi.org/10.3390/jtaer16070145>
- Wahid, R., Karjaluo, H., & Taiminen, K. (2022). How to Engage Customers on TikTok? ICIS 2022 Proceedings. <https://aisel.aisnet.org/icis2022/social/social/11>
- Wang, J. (2025). Quantitative analysis of short video dissemination effects: Based on the likes, share rate and comment density: *Proceedings of the 2nd International Conference on Innovations in Applied Mathematics, Physics, and Astronomy*, 360–363. <https://doi.org/10.5220/0013825800004708>
- Wang, Y. (2021). *Multimodal analysis: Researching short-form videos and the theatrical practices*. <http://hdl.handle.net/10125/76003>
- Watier, N. (2023). Subtle threat cues in marketing horror and children's entertainment. *Psychology of Popular Media*, 12(2), 231–241. <https://doi.org/10.1037/ppm0000401>
- Wickham, H. (2009). *Stringr: Simple, consistent wrappers for common string operations* [Data set]. <https://doi.org/10.32614/CRAN.package.stringr>
- Wood, S. N. (2017). *Generalized additive models: An introduction with r* (2nd edn). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315370279>
- Wood, S., & Scheipl, F. (2009). *Gamm4: Generalized additive mixed models using 'mgcv' and 'lme4'* [Data set]. <https://doi.org/10.32614/CRAN.package.gamm4>
- Xiao, L., Li, X., & Zhang, Y. (2023). Exploring the factors influencing consumer engagement behavior regarding short-form video advertising: A big data perspective. *Journal of Retailing and Consumer Services*, 70, 103170. <https://doi.org/10.1016/j.jretconser.2022.103170>
- Xiao, Q., Huang, W., Qu, L., & Li, X. (2025). The impact of multimodal information features of short sales videos on consumer engagement behavior: A multi-method approach. *Journal of Retailing and Consumer Services*, 82, 104136. <https://doi.org/10.1016/j.jretconser.2024.104136>

- Yadavalli, A., Pimentel, T., Regev, T. I., Wilcox, E., & Warstadt, A. (2025). *What do prosody and text convey? Characterizing how meaningful information is distributed across multiple channels*. arXiv.
<https://doi.org/10.48550/ARXIV.2512.16832>
- Yang, E. F., Kriss, L. A., & Sun, Y. (2024). Fun with frustration? Tiktok influencers' emotional expression predicts user engagement with covid-19 vaccination messages. *Health Communication*, 39(12), 2458–2473. <https://doi.org/10.1080/10410236.2023.2259621>
- Zannettou, S., Nemes-Nemeth, O., Ayalon, O., Goetzen, A., Gummadi, K. P., Redmiles, E. M., & Roesner, F. (2024). Analyzing User Engagement with TikTok's Short Format Video Recommendations using Data Donations. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–16.
<https://doi.org/10.1145/3613904.3642433>
- Zeng, J., Abidin, C., & Schäfer, M. S. (2021). Research perspectives on TikTok and its legacy apps: introduction. *International Journal of Communication*, 15, 3161–3172.
<https://ijoc.org/index.php/ijoc/article/view/14539/3494>
- Zhang, J., Synave, R., Delepoulle, S., & Cozot, R. (2023). Reconstructing image composition: Computation of leading lines. *Journal of Imaging*, 10(1), 5. <https://doi.org/10.3390/jimaging10010005>
- Zhang, Y., Han, M., Swerts, M., & Gu, Y. (2025). The effect of gender on multimodal child-directed language: Evidence from analyses of broadcast programmes. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 47(0). <https://escholarship.org/uc/item/86v6q19b>

Appendices (Chapter 2):

1. Sampling procedure:

Figure 21 Sampling (TikTok and Douyin)



2. Data descriptions

Table 9 Total number of videos of selected content creators

Username handle	Gender	Nationality	Video count (Research API)	Video count (Analisa.io)	Video count (Manual Calibration)	No. of transcriptions (Research API)	No. of transcriptions (third-party services)	Total number of transcripts
lizzza	female	American	159	N/A	168	15	3	18
daniel. labelle	male	American	302	319	319	19	0	19
hannahstocking	female	American	606	N/A	691	25	5	30
jeremyllynch	male	British	792	738	734	91	9	100
adamw	male	American	632	567	584	106	37	143
spencewuah	male	American	2,346	457	449	162	1	163
adrianbliss	male	British	106	69	113	0	12	12
steven_he	male	American (Chinese Irish actor)	199	237	237	35	0	35
hayleygeorgiamorris	female	British	187	204	201	0	20	20
jackjos3ph	male	British	832	806	1,028	0	41	41
nicholas_flannery	male	British	842	902	890	13	0	13
ameliadimz	female	British	141	N/A	156	21	1	22
itscaitlinhello	female	American	160	170	170	52	0	52
maddiegracejepson	female	British	580	N/A	592	108	12	120
Total Number of Transcripts						647	141	788
Number of Participants			14	10	14			
Number of Videos			7,884	4,469	6,373			
Engagement Metrics			view, likes, comments, shares and transcriptions	view, likes, comments, shares	view, likes, comments, shares			

Time of collection

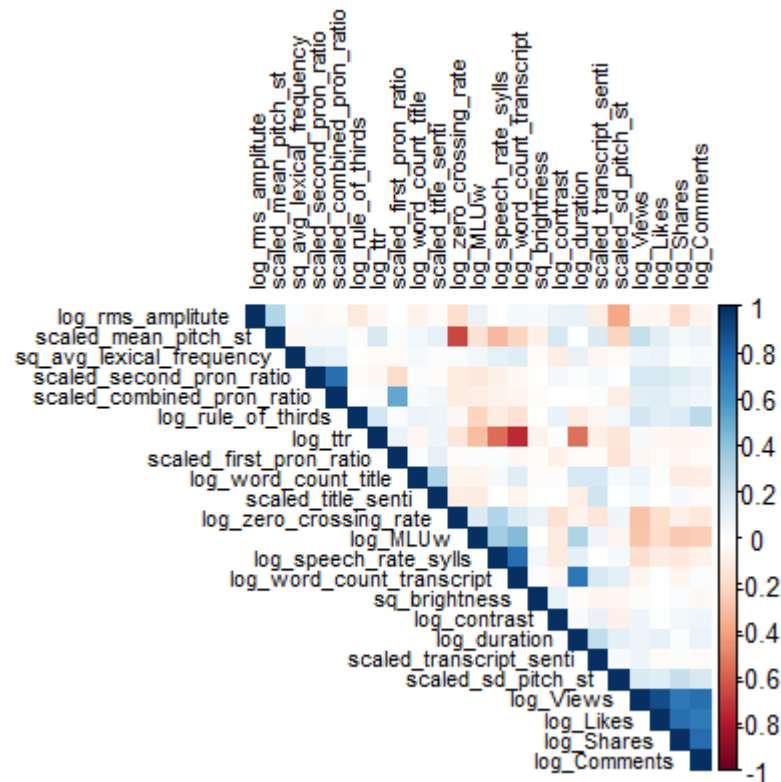
Apr-24

Mar-23

Between April
and July 2023

Note: The content creators analysed in this study have not been anonymized, as they are public figures (Stevens et al., 2015). This approach parallels established conventions in SM research, where participant identities are likewise reported without pseudonymization (e.g., Alexandre et al., 2022; Arora et al., 2019).

3. Correlation



Note: Correlations were generally low to moderate ($|r| < .55$), indicating limited multicollinearity.

4. Interaction terms:

Categorical-Categorical variables

```
> results %>% arrange(desc(cramers_v))
# A tibble: 30 × 4
  var1      var2      chi_p crammers_v
<chr>    <chr>    <dbl>    <dbl>
1 filter  sound_effect 0.000000146 0.303
2 sound_effect filter 0.000000146 0.303
3 gender  nationality 0.0001000 0.275
4 nationality gender 0.0001000 0.275
5 filter  close_ties 0.000000272 0.205
6 close_ties filter 0.000000272 0.205
7 sound_effect close_ties 0.000999 0.135
8 close_ties sound_effect 0.000999 0.135
9 gender  stitch_duet 0.000105 0.134
10 stitch_duet gender 0.000105 0.134
```

Categorical-continuous variables:



5.

```
> ##GAMM model (likes): separate linear smomoth terms
Family: gaussian
Link function: identity

Formula:
log_Likes ~ log_views + log_duration + log_rule_of_thirds + sq_brightness +
  log_rms_amplitude + s(scaled_sd_pitch_st, by = gender, k = 5) +
  s(log_ttr, k = 5) + s(log_word_count_title, k = 10) + s(scaled_transcript_senti,
k = 5) + s(scaled_title_senti, by = gender, k = 5) + filter

Parametric coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.42816    0.36969  -6.568 9.77e-11 ***
log_views      1.00683    0.01538  65.462 < 2e-16 ***
log_duration   0.05708    0.02448   2.332  0.01997 *
log_rule_of_thirds 0.06292    0.02107   2.986  0.00292 **
sq_brightness  -0.04086    0.01814  -2.253  0.02457 *
log_rms_amplitude -0.06426    0.02246  -2.862  0.00434 **
filterYes     -0.16957    0.07888  -2.150  0.03191 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
              edf Ref.df    F p-value
s(scaled_sd_pitch_st):genderfemale 3.068 3.068 10.239 1.5e-06 ***
s(scaled_sd_pitch_st):gendermale   1.342 1.342  0.151  0.6903
s(log_ttr)                         2.059 2.059  3.532  0.0346 *
s(log_word_count_title)            3.397 3.397 14.078 < 2e-16 ***
s(scaled_transcript_senti)         1.674 1.674  3.558  0.0777 .
s(scaled_title_senti):genderfemale 2.682 2.682  3.043  0.0250 *
s(scaled_title_senti):gendermale   1.000 1.000  0.612  0.4342
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.763
lmer.REML = 1144.2 Scale est. = 0.2245    n = 741
> AIC(gam_mixed_likes_6$lmer)
[1] 1190.233
```

```
##GAMM model (comments): separate linear smmoth terms
> summary(gam_mixed_comments_6$gam)

Family: gaussian
Link function: identity

Formula:
log_Comments ~ log_Likes + s(log_duration, k = 10) + s(scaled_sd_pitch_st,
  by = gender, k = 5) + s(log_ttr, by = gender, k = 5) +
log_word_count_title +
  s(scaled_transcript_senti, by = gender, k = 5) + s(scaled_title_senti,
  k = 5) + s(log_MLUw, by = gender, k = 5) + filter

Parametric coefficients:

```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-4.76580	0.69776	-6.830	1.81e-11	***
log_Likes	1.01072	0.01877	53.844	< 2e-16	***
log_word_count_title	-0.10372	0.02818	-3.681	0.00025	***
filteryes	-0.26521	0.10965	-2.419	0.01582	*

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:

```

	edf	Ref.df	F	p-value	
s(log_duration)	2.865	2.865	7.209	0.000102	***
s(scaled_sd_pitch_st):genderfemale	1.000	1.000	0.037	0.847275	
s(scaled_sd_pitch_st):gendermale	1.000	1.000	10.980	0.000967	***
s(log_ttr):genderfemale	2.852	2.852	5.256	0.001720	**
s(log_ttr):gendermale	1.000	1.000	1.316	0.251689	
s(scaled_transcript_senti):genderfemale	3.082	3.082	6.636	0.000204	***
s(scaled_transcript_senti):gendermale	1.436	1.436	0.315	0.526147	
s(scaled_title_senti)	3.317	3.317	3.567	0.030795	*
s(log_MLUw):genderfemale	2.254	2.254	4.762	0.007734	**
s(log_MLUw):gendermale	1.000	1.000	0.14	0.906446	

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.469
lmer.REML = 1605.5  scale est. = 0.41445  n = 741
> AIC(gam_mixed_comments_6$mer)
[1] 1657.488

```

```
##GAMM model (comments): separate linear smomoth terms
> summary(gam_mixed_shares_6$gam)

Family: gaussian
Link function: identity

Formula:
log_Shares ~ log_Likes + log_duration + s(log_contrast, k = 10) +
  log_speech_rate_sylls + s(log_ttr, by = gender, k = 5) +
  s(log_word_count_title, k = 10) + scaled_transcript_senti +
  nationality + close_ties

Parametric coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -7.44592    0.55172  -13.496 < 2e-16 ***
log_Likes       1.19802    0.02352   50.938 < 2e-16 ***
log_duration   -0.08751    0.04239   -2.065  0.03932 *
log_speech_rate_sylls -0.12123    0.04604   -2.633  0.00864 **
scaled_transcript_senti -0.07303    0.03551   -2.056  0.04009 *
nationalityBritish  1.26721    0.62798    2.018  0.04397 *
nationalityChinese-Irish -0.06154    1.21613   -0.051  0.95965
close_tiesYes    0.31170    0.07414    4.204 2.95e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
              edf Ref.df      F  p-value
s(log_contrast)    1.875   1.875   4.979 0.015874 *
s(log_ttr):genderfemale 1.000   1.000   0.012 0.911608
s(log_ttr):gendermale  1.000   1.000  12.654 0.000399 ***
s(log_word_count_title) 2.453   2.453  11.629 4.06e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.645
lmer.REML = 1909.3  scale est. = 0.68284    n = 741
> AIC(gam_mixed_shares_6$mer)
[1] 1945.268
```

Chapter 4: Character Performance and Engagement: Multimodal Production Elements and Gender Effects in TikTok Role-Play Comedy

Abstract

Role-play comedy on TikTok, in which creators embody characters distinct from their everyday personas, is both prevalent and poorly understood. How such content is constructed multimodally, and whether it independently predicts audience engagement, remains empirically unresolved. The study employs mixed-effects and generalised additive mixed models to analyse 741 comedic videos from 14 high-profile creators active during 2020 to 2022.

Results demonstrate that role-play videos exhibit a distinctive multimodal production signature, including more accessories, distinct clothing, relationship-focused narratives, and characteristic audio and verbal patterns. Role-play predicted approximately 21 per cent higher likes but showed no significant effect on comments. This asymmetry reveals engagement as multidimensional rather than unitary. A gender interaction shaped shares: male creators derived substantial gains from same-gender role-play despite showing lower baseline sharing, whilst female creators showed minimal role-play effects. These patterns suggest that character embodiment constitutes gender-atypical behaviour for male creators, enhancing novelty and appeal.

The findings reveal that TikTok role-play is a distinct creative mode, operating through theatrical character construction rather than everyday persona management. Engagement, this study suggests, is multidimensional: passive approval, discursive participation, and content amplification respond differently to performance strategies. Practically, role-play represents an accessible strategy for enhancing likes and shares without requiring elaborate production infrastructure. In Conclusion, platform designers should account for these differential effects when calibrating recommendation algorithms.

Keywords: TikTok, role-play, engagement metrics, multimodal production elements, gender performance, short-form video, comedic performance, digital platforms

Introduction

Short-form video platforms have fundamentally reshaped the landscape of digital communication and entertainment. TikTok, the international version of China's Douyin, surpassed 2 billion downloads by April 2020 and sustained remarkable growth throughout the COVID-19 pandemic (Kaye et al., 2021; Qu, 2020). Its algorithmic architecture prioritises engagement metrics, generating feedback loops that reward high-performing content (Boeker & Urman, 2022; Guinaudeau et al., 2022). Creators must therefore design content that captures attention rapidly and encourages active viewer participation (Xiao et al., 2025). Typical video lengths of 30 to 60 seconds intensify demands on performance, compelling creators to deploy visual and auditory cues efficiently (Xiao et al., 2023). Comedic strategies have proliferated in this context as a primary mode of creative engagement.

Creators on TikTok employ a diverse range of comedic strategies to attract and retain audiences. These include lip-syncing, duet and stitch interactions, and point-of-view skits, as well as parody, irony, and wordplay (Cervi

& Divon, 2023; Divon & Eriksson Krutrök, 2024; Matamoros-Fernández, 2023; Zeng & Abidin, 2021). Such strategies reflect the platform's affordances for highly imitable, rapidly iterated performance. Role-play stands out as one of the most prevalent of these approaches, yet it remains among the least systematically examined. More specifically, role-play involves the explicit adoption of distinct characters to enact scenarios, dialogues, or narratives through coordinated multimodal cues (Cooper, 2023). Unlike the ambient self-presentation characteristic of most social media activity, role-play renders performance overt and theatrical. It draws upon voice modulation, costume changes, editing cuts, and stage properties, often enabling a single creator to portray multiple characters (Xing, 2024). This multimodal richness is central to role-play's appeal, yet how such content is constructed and received on TikTok remains poorly understood. Whether role-play operates as an independent driver of engagement, or merely signals high-investment content more generally, cannot be determined from existing accounts.

Audience engagement research has identified humour, novelty, emotional resonance, and production quality as significant predictors across platforms (Dong et al., 2024; Lu & Shen, 2023; Xie et al., 2025). Studies on multimodal features demonstrate that visual composition, audio characteristics, and linguistic markers each independently shape viewer responses (Cheng & Li, 2024; Yang et al., 2025). Yet few investigations have considered these features in combination with broader performance strategies; most examine one or two cues in isolation. Role-play, which combines these features into a coherent performance format, is therefore particularly difficult to examine through single-cue approaches.

Role-play typically integrates humour, vocal performance, and material production to generate entertainment and interactivity (Ao et al., 2023). Qualitative documentation of this practice exists (Xing, 2024), but quantitative analyses isolating its engagement effects whilst controlling for correlated production characteristics remain largely absent. It therefore remains unclear whether role-play independently drives engagement or merely correlates with high-investment content. This uncertainty has consequences for creator strategy and platform design alike.

Gender performance within role-play raises pressing empirical questions. Cross-gender portrayal, whereby creators embody characters of a different gender, constitutes a distinctive subset of role-play practice.

Theoretical frameworks rooted in Butler's gender performativity and Goffman's dramaturgical approach suggest that gender boundary crossing may generate engagement through incongruity and subverted expectations (Barta & Andalibi, 2021; Gran, 2025). Research documents a diversification from male-to-female toward female-to-male and dual-role enactments on TikTok (Xing, 2024), whilst gender-atypical content elsewhere generates higher engagement through expectation violation (Wu et al., 2025). Cross-gender portrayal involves distinct vocal, costuming, and gestural cues (Morgenroth & Ryan, 2021; Perugia & Lisy, 2023; Tang et al., 2021; Xu et al., 2024), which have been linked to differential audience engagement in related contexts (Åkestam et al., 2021; Wu et al., 2025). Whether cross-gender role-play shapes likes, comments, and shares differently, and how creator gender moderates these effects, represents a significant open question.

Role-play encompasses diverse character configurations beyond gender variation, including perspective-taking across generational positions within a single scenario (Han & Ho, 2024; Parvin, 2024). Parenting humour and parasocial family dynamics constitute a prevalent subset within relationship-focused comedy (Carter & Davies, 2025; Osisanwo et al., 2024). These intergenerational scenarios depict parent-child, grandparent-grandchild, or mentor-mentee dynamics, enabling the dramatisation of generational tensions and family communication patterns (Smith & Mendelson, 2025). Such performances may generate distinct audience responses by invoking personal family experiences or prompting reflection on generational differences. Their effects on engagement, however, remain unexamined.

Two persistent methodological problems complicate the interpretation of existing findings. Most studies treat engagement as unidimensional, yet the processes underlying likes, comments, and shares are qualitatively distinct (Habibi & Salim, 2021; Shahbaznezhad et al., 2021). Whether role-play differentially predicts these outcomes has not been examined, obscuring distinct motivations underlying different engagement behaviours. The nested structure of videos within creators poses a second problem. Content from the same creator is not independent, owing to consistent styles, recurring audiences, and varying baseline popularity. Failing to account for this hierarchy through multilevel modelling can produce misleading estimates of content characteristic effects (Barrie et al., 2025).

The present study directly addresses both problems through systematic quantitative analysis. First, it identifies which production elements distinguish role-play from non-role-play content, accounting for creator-level variation. Second, it examines how production elements and role-play characteristics predict audience engagement, employing generalised additive mixed models to capture non-linear relationships and creator effects. This approach illuminates not only what characterises role-play content but also the mechanisms through which such content achieves engagement. These questions are particularly timely given TikTok's algorithmic prioritisation of engagement metrics and its continuing influence on creator practice.

This research makes several contributions to ongoing discussions about creative labour in platform economies. Theoretically, it extends performance theory into short-form digital contexts whilst advancing understanding of engagement as a multidimensional construct. It further provides the first quantitative examination of cross-gender role-play effects on multiple engagement metrics in algorithmically mediated environments. Empirically, it offers rigorous quantitative estimates of role-play effects, disentangling performance strategies from correlated production characteristics. Practically, it informs content strategy decisions by clarifying which elements independently drive different forms of audience response.

Literature Review

2.1 Engagement on Short-Form Video Platforms

Short-form video platforms, particularly TikTok, have transformed digital entertainment consumption and creator-audience dynamics. Understanding engagement requires examining multiple dimensions: how it is conceptualised, what content characteristics drive responses, how creator attributes shape reception, and how algorithms amplify reach (Aldlimi et al., 2025; Bouchaud, 2024; Jaakonmäki et al., 2017; Trunfio & Rossi, 2021).

Table 1 synthesises key findings across these dimensions, providing foundational context for the present study's focus on role-play and gender performance. The evidence establishes that engagement emerges from complex interactions between content features, creator characteristics, and platform affordances. Yet most studies examine these elements in isolation or focus on general content rather than specific performance strategies. The

following sections examine role-play and gender performance, which remain underexplored despite their prevalence on TikTok.

Table 10 Engagement on Short-Form Video Platforms: Key Concepts and Determinants

Dimension	Key Findings	Supporting Studies
<i>Conceptualising Engagement</i>		
Engagement metrics	Likes, views, and shares serve as proxies for user interest but may obscure behavioural complexities; likes represent low-effort acceptance; comments demand greater investment and facilitate discussion; sharing indicates active endorsement and distribution	Achen et al., 2024; Aldlimi et al., 2025; Habibi & Salim, 2021; Kuo & Chen, 2025; Li et al., 2021; Shahbaznezhad et al., 2021; Tenenboim, 2022; Yu et al., 2025
Alternative measures	Engagement-based analyses may misrepresent attention dynamics; attention-based measures (watch time, replays, scrolling pauses) provide fuller picture; engagement measures attention that is inherently fluid and unstable	Barrie et al., 2025; Li et al., 2021; Liu et al., 2024; Zhao et al., 2025
<i>Content Characteristics</i>		
Humour and emotion	Humour consistently drives engagement; playful activism and gallows humour make content shareable; empathy and relevance foster parasocial connections; both positive and negative emotions increase engagement; mixed emotional trajectories sustain engagement longer	Lu & Shen, 2023; Xie et al., 2025; Cervi & Divon, 2023; Hautea et al., 2021; Herrick et al., 2021; Zhao et al., 2025; Dong et al., 2024; Munaro et al., 2021; Li et al., 2021; Rejer et al., 2024
Novelty and information	Innovative topics attract engagement regardless of subscriber base; informational content outperforms purely emotional posts on TikTok	Hussain et al., 2024; Wahid et al., 2023
Production quality	Close-up shots and second-person views predict higher engagement; brightness, face presence, and meme visuals correlate with increased engagement; platform features (duets, stitches) encourage collaborative engagement; shorter videos (under 60 seconds) gain more engagement	Ahlse et al., 2020; Cheng & Li, 2024; Lu & Shen, 2023; Ngilangil, 2022
Multimodal integration	Greater visual variation (shot changes, motion, scene variety) shows positive relationship with interaction; moderate emotional arousal in voice and music maximises engagement; alignment between sound arousal and visual variation boosts interaction	Yang et al., 2025; Zhang et al., 2022
Visual and linguistic features	Human characters, subtitles, and clear messaging enhance sharing; visual aesthetics shape interpretation; sentiment and second-person pronouns affect engagement; authenticity and relatability vital for engagement	Chen et al., 2021; Cheng & Li, 2024; Heikkonen, 2024; Katiaj et al., 2024; Li et al., 2021; Wang, 2021;
Theoretical measures	Visual composition features (brightness, contrast, rule of thirds) direct attention; lexical diversity and utterance length indicate communicative complexity; RMS amplitude aligns with arousal-attention-memory models	Chen et al., 2024; Kontogiorgos & Gustafson, 2021; Kostyk & Huhmann, 2021; Lee et al., 2024; Shen & Pritchard, 2022; Mather & Sutherland, 2011; Simonetti & Bigne, 2022
<i>Creator Characteristics</i>		
Demographics and authenticity	Age, gender, and nationality shape content style and reception; modest followings can generate viral content; how creators express themselves (emojis, emotions) more influential than demographic characteristics; consistent self-presentation builds authenticity	Agrawal, 2024; Bhandari & Bimo, 2022; Meng & Leung, 2021; Zhao et al., 2025; Barta & Andalibi, 2021
Influencer attributes	Micro-influencers perceived as more authentic; credibility attributes (expertise, trustworthiness, similarity, attractiveness) correlate with engagement; parasocial relationships increase engagement outcomes; congruence between influencer and follower amplifies effects	Ashraf et al., 2023; Chan et al., 2023; Ao et al., 2023; Pan et al., 2025; Venciute et al., 2023
<i>Algorithmic Amplification</i>		
Algorithm mechanics	TikTok's algorithm prioritises engagement signals over follower networks; perceived recommendation accuracy, serendipity, and effortlessness increase flow experience; early engagement indicators trigger broader distribution creating feedback loops	Boeker & Urman, 2022; Chen, 2024; Guinaudeau et al., 2022; Klug, 2020; Su et al., 2021; Xiao et al., 2025; Yin et al., 2024;
User behaviour	Users experience lean-back, television-like stream where system chooses next clip; algorithm clusters users into hyper-specific niches creating sense of belonging; users develop algorithmic engagement to train algorithm and maintain preferred feed; folk theories about algorithm guide behaviour	Bhandari & Bimo, 2022; Firth & Marinelli, 2025; Schellewald, 2022; Simpson et al., 2022

2.2 Role-Play and Performance in Digital Media

Role-playing can increase audience engagement when it adds entertainment, emotion, and interactivity, but its impact depends strongly on content context and platform. Its most direct pathway to engagement is entertainment value: humour, storytelling, and character-driven formats reliably generate more interactions than purely informational posts (Ao et al., 2023; Vrontis et al., 2021). These effects are nonetheless context-dependent, varying by platform and campaign type. Role-play content is inherently multimodal, combining voice modulation, costuming, stage properties, and editing to construct character. This richness distinguishes it from the ambient self-presentation characteristic of most social media activity.

Beyond entertainment, role-play generates perceptual, social, and embodied experiences that facilitate sharing, commenting, and co-developing content (Pradhan et al., 2023). For example, parasocial interaction (one-sided relational bond), identification, and perceived similarity mediate these outcomes, as viewers develop connections with recurring characters or personas (Pradhan et al., 2023; Vrontis et al., 2021). Whether these bonds strengthen or erode depends on authenticity: relatable, genuine characters deepen connection (Ballester et al., 2025), whilst forced or overly commercial acting reduces trust (Martínez-López et al., 2020; Pfeuffer & Huh, 2021). The emotional tone of role-play makes its effectiveness context-dependent, with emotional content outperforming rational content in some platform contexts but not others (Shahbaznezhad et al., 2021).

Yet the evidence base has significant analytical shortcomings. Most research on role-play in digital contexts has been qualitative or descriptive, documenting its prevalence without rigorously testing its effects on engagement. How role-play content is multimodally constructed on TikTok remains undocumented, and whether its constituent features independently drive engagement has not been established. The few quantitative studies examining character-driven content tend to conflate role-play with correlated variables such as humour or production quality (Shahbaznezhad et al., 2021). This conflation makes it difficult to isolate role-play's independent contribution or assess its weight relative to other features affecting audience metrics. Whether role-play effects vary across engagement types is similarly unexamined: role-play might enhance passive approval without necessarily stimulating commenting or content amplification. This unidimensional treatment of engagement obscures the distinct motivations underlying different audience behaviours.

2.3 Gender Performance and Cross-Gender Portrayal

2.3.1 Gender as Performance

Scholarly accounts of gender performance draw on two foundational but distinct frameworks. Butler's (1990) performativity theory holds that gender is not a fixed attribute of persons. Rather, it is constituted through the iterative citation of cultural norms, enacted in repeated everyday behaviour. Butler's concern is with unreflective, mundane enactment rather than deliberate theatrical staging. Gender accrues as an effect of accumulated acts rather than expressing a prior identity. Goffman's (1959) dramaturgical framework approaches everyday social interaction differently, treating it as impression management performed before an audience. Both frameworks were originally developed to explain everyday self-presentation rather than explicitly staged performance.

Explicitly theatrical performance of the kind found in TikTok role-play differs from the everyday gender enactment these frameworks describe. In role-play, creators consciously adopt characters, and audiences understand the performance as such. This reflexivity is significant. Butler's (1990) account of gender as citational practice nonetheless remains relevant: cross-gender role-play quotes and potentially destabilises normative gender presentations precisely by making the citational structure visible. Similarly, Goffman's (1959) concept of front-stage performance provides analytical purchase on how creators manage impressions and frame performances for audiences. Recent scholarship has extended both frameworks to explicitly theatrical digital performance (Gran, 2025), confirming their value even where performance is overtly staged.

Research across performance traditions illuminates how gender is constructed through coordinated multimodal cues rather than any single channel. In the Japanese comic tradition of Rakugo, a single performer shifts voice, speech style, and bodily stance to portray characters of different genders (Şahin, 2022). The conventional unsuitability of a female performer's voice in this genre reveals how tightly vocal production, gender ideology, and comic convention are bound together. Social-semiotic analyses of stand-up impersonation demonstrate that comedians deploy phonology and voice quality to embody characters and manage audience rapport, whilst retaining a distinct authorial persona (Logi, 2025). Experimental work on acting confirms that convincing character portrayal requires coordinated multichannel reinforcement: body, face, pitch, and loudness must shift

together (Berry et al., 2022). Gender performance, on this account, cannot be reduced to isolated production choices.

Research on short-form video platforms provides additional evidence of the contested nature of gender performance in digital contexts. Li and Wu (2025) analysed male beauty vloggers on Bilibili, finding that creators construct resistance identities through dramatic makeup and cross-gender presentation, constrained by platform censorship. These findings illuminate the interplay between overt performative choices, including costume, makeup, and voice modulation, and the constraints imposed by algorithmic and regulatory systems. The study is confined to a single platform and a specialised category; it does not examine comedic role-play or quantify engagement effects. Whether similar patterns characterise TikTok role-play creators, and whether these production choices differentially predict audience engagement metrics, remains an open empirical question. Cross-gender performance in older traditions provides further conceptual resources. Ethnographic accounts of Cantonese opera cross-dressing foreground androgynous embodiment through stylised posture, gesture, codified costuming, and carefully managed offstage persona (Tse, 2021). The aim is not to pass as another gender but to sustain a productive ambiguity that audiences read as emotionally charged. Historical analysis of early cinema cross-dressing demonstrates how costume, wig, and gestural affect generate ambiguity that may be read as comic, glamorous, or transgressive (O'Rourke, 2020). These traditions underline that cross-gender performance operates through recognisable conventions that simultaneously cite and potentially destabilise normative gender presentations.

Analyses of women's stand-up comedy offer coding schemes for gendered self-presentation, body display, and vocal affect adaptable to cross-gender role-play research (Cook et al., 2024; Crigler, 2024; Li & Guo, 2025; Lockyer & De Benedictis, 2023). Linguistic work on dialect and enregistered speech styles as comedic resources offers complementary tools for examining gendered vocal deployment on TikTok (Gerwin, 2024; Puterbaugh & Medford, 2025). Collectively, these bodies of work demonstrate that gender performance is a richly multimodal and convention-bound practice. Its specific mechanisms on short-form video platforms, and their quantifiable consequences for engagement, remain inadequately examined.

Shared cultural assumptions about gender set the interpretive framework through which audiences evaluate performance. These assumptions establish normative expectations about masculine and feminine behaviour, speech, and self-presentation (Charlesworth et al., 2021; Eagly & Karau, 2002; Morgenroth & Ryan, 2021). Content that meets those expectations tends to be processed with minimal resistance. Departures are a different matter. They may attract heightened attention and generate novelty-driven engagement, but they can equally trigger resistance or reduced algorithmic amplification. The direction of the effect is context-dependent. Comedic genre is distinctive in the cultural latitude it grants for norm transgression. A comedic generic frame signals to viewers that exaggeration, inversion, and gender-norm violation are anticipated and welcomed rather than aberrant. Comedy audiences do not merely accommodate such boundary crossings. They seek them out as a condition of comedic pleasure (Kuipers, 2002). This is particularly consequential for gendered performance in platform contexts.

The comedic frame creates a sanctioned space for gender-transgressive performance that everyday interaction does not easily afford. When a male creator performs stereotypically feminine mannerisms, the comedic frame renders that departure legible as entertainment rather than violation. So too when a female creator adopts hyper-masculine vocal patterns and behaviour. Whether viewers respond positively or negatively to such performances, and whether the response differs by the performer's own gender, remains empirically unresolved. These considerations underpin the decision to model creator gender as a structural moderator in the present study. Gender is not treated here as a background covariate. It is a theoretically grounded analytical dimension along which audience response to role-play is expected to vary. The cultural consequences of norm transgression differ depending on who is doing the transgressing.

2.3.2 Cross-Gender Role-Play on Social Media

Cross-gender role-play, whereby creators explicitly embody characters of a different gender (also referred to as cross-gender portrayal), constitutes a prevalent but under-researched content category on TikTok. These performances range from subtle adoptions of vocal qualities associated with another gender to elaborate productions involving full costume, prosthetic makeup, and exaggerated affectation. Much cross-gender content occurs within multi-character sketches, in which a single creator portrays characters of contrasting genders to

enact interpersonal scenarios. The empirical evidence base, however, remains emergent and multidisciplinary, with TikTok-specific quantitative research particularly sparse.

The most substantive TikTok-specific analysis is Xing's (2024) qualitative study of 424 gender-emotion videos. It documented a diversification from male-to-female impersonation toward female-to-male and dual-role enactments, alongside increased aestheticisation and the influence of couple culture. Xing argues that these practices partially deconstruct gender binaries without overturning entrenched norms. Critical discourse analysis of Douyin performers similarly reveals reliance on recurrent gestural, vocal, and narrative tropes to embody stereotyped femininity (Tang et al., 2021). Portrayals of women as fragile or insincere within domestic scenarios reinforce traditional gender ideology, even whilst playing with its conventions.

Both studies are valuable, but descriptive in orientation. They document what cross-gender performances look like rather than measuring their effects on audience engagement. Neither systematically analyses the production features through which characters are constructed, nor quantifies whether those features predict likes, comments, or shares. The present study addresses both problems by combining systematic multimodal measurement with engagement modelling.

A small body of related work examines adjacent practices on TikTok and other platforms. A phenomenological study of Muslim adolescents found ambivalence toward gender-imitation content on TikTok, Instagram, and YouTube, with entertainment value weighed against moral and religious judgements (Suhertina et al., 2025). Its small sample of 15 participants and focus on spectatorship rather than production, however, leave engagement effects unaddressed. Skorodihina (2024) finds that short-form affordances facilitate identity exploration through performative fan labour in Genshin Impact communities, though the fandom context limits generalisability. Brush (2024) and Ben-Iheanacho (2024) enrich the theoretical vocabulary of cross-gender performance online, yet neither provides quantitative evidence on engagement outcomes. In sum, this body of work establishes the cultural significance and prevalence of cross-gender practices without resolving how they function as drivers of audience engagement.

Evidence from research on gender-atypical content more broadly offers indirect but relevant insights. Wu et al. (2025) found that executives on microblogs receive greater engagement when posting gender-incongruent

content, a pattern attributed to expectation violation. Female executives posting work-focused content and male executives posting life-focused content each attracted higher engagement than gender-congruent posts. In influencer marketing, gender dynamics operate differently: women display stronger parasocial affinity with female influencers, boosting engagement, whilst men show little systematic gender preference (Hudders & De Jans, 2022). On YouTube, content attracting gender-balanced audiences achieves greater long-term reach than content skewed toward a single gender (Thelwall & Foster, 2021). These findings suggest that whilst gender incongruence may generate short-term engagement through novelty, sustained reach may depend on broader audience appeal. Whether these patterns hold for TikTok role-play remains untested.

Creators across platforms also deploy gender performance strategically to manage audience expectations and platform affordances. Entrepreneurs on Twitter (X) of both genders blend masculine and feminine traits, including emotional expressiveness and gratitude, to navigate the double bind between credibility and likeability (Liu et al., 2025). Virtual YouTubers (VTubers) frequently adopt highly stylised gender expressions that cross or exaggerate normative gender boundaries (Wan & Lu, 2024). This can foster more inclusive chat environments and reduce visible harassment, though it raises concerns about self-objectification. These examples suggest that strategic gender performance is understood by creators as a lever for engagement, though the mechanisms through which it operates remain underspecified.

Structural factors also shape the reach of cross-gender content. Recommendation algorithms have been shown to reproduce binary and stereotypical gender roles, constraining the visibility of non-conforming performances (Fosch-Villaronga et al., 2021; Gerrard & Thornham, 2020). This algorithmic bias may limit cross-gender role-play's reach independently of its intrinsic appeal. Cross-gender performance can additionally expose creators, particularly women and gender minorities, to heightened scrutiny and harassment, potentially deterring such content even where engagement benefits exist.

2.4 Research Gaps and Research Questions

2.4.1 Analytical Challenges

Four analytical problems motivate the present study's design. First, role-play's multimodal production characteristics on TikTok have not been systematically described, and its independent effects on engagement remain unestablished. Second, how cross-gender role-play is multimodally represented, and whether such

representations differentially predict engagement compared to same-gender role-play, remains empirically unresolved. Third, engagement has predominantly been treated as a unidimensional construct, yet likes, comments, and shares reflect qualitatively distinct processes whose differential drivers have not been examined. Fourth, as the foregoing review demonstrates, the nested structure of videos within creators has been inadequately addressed, risking misleading estimates of content characteristic effects (Barrie et al., 2025). Taken together, these problems call for comprehensive multimodal measurement, appropriate hierarchical modelling, and treatment of engagement as a multidimensional outcome.

2.4.2 Research Questions

The present study investigates two primary questions:

***RQ1:** Which production elements, including costume design, makeup and hairstyling, visual design, audio characteristics, and screenplay features, significantly differ between role-play and non-role-play comedic videos on TikTok, accounting for creator-level variation?*

***RQ2:** How do production elements and role-play features influence audience engagement, including likes, comments, and shares, in short-form comedic TikTok videos, whilst controlling for creator effects and engagement interdependencies?*

***RQ2a:** Does role-play presence significantly predict engagement metrics when controlling for all production elements?*

***RQ2b:** How do cross-gender role-play performances relate to audience engagement metrics?*

Addressing these questions contributes theoretically by clarifying how performance characteristics influence engagement, and empirically by providing rigorous quantitative estimates of role-play effects. Practically, findings inform creator content strategy decisions by clarifying which elements independently drive different forms of audience response. The research extends performance theory into short-form digital contexts whilst advancing understanding of engagement as a multidimensional construct responsive to content, creator, and audience factors.

Methodology

3.1 Participants and Procedures

Sample Composition

The participant pool comprised 14 high-profile comedy creators active on TikTok between 1 January 2020 and 31 December 2022. This temporal window captured the acute pandemic phase and subsequent recovery,

enabling examination of comedic engagement across distinct sociocultural moments (Davidson et al., 2023). TikTok comedy relies on vivid character-driven performance shown to engage attention and reflect lived experience (Nishimoto et al., 2011; Pearson et al., 2015, 2019). Prominent comedic creators were therefore selected as a purposive sample for examining societal narratives during this period. Creators were identified through Google search visibility and viral content indicators, ensuring coverage of influential voices. The sample included six UK-based creators (three male, three female) and six US-based creators (three male, three female).

Two accounts were treated as optional: @steven_he, who identified as Chinese-Irish whilst residing in the United States, and @nicholas_flannery, who began uploading only in 2021. Detailed sampling procedures appear in Appendix 1.

Inclusion Criteria and Final Dataset

Video selection applied two criteria: transcription availability and inclusion within each creator's top ten per cent of most-viewed content during the study period. Transcription data were obtained from the TikTok Research API or third-party speech-to-text services. Videos featuring purely visual humour without verbal content were excluded. Such videos lack transcript data and cannot provide the multimodal coverage required for analysis across visual, audio, and linguistic dimensions.

The 14 creators published 6,373 videos collectively. Of these, 788 possessed complete transcriptions and met inclusion criteria. A further 47 proved unavailable during data collection in September 2024, rendered inaccessible by UK age-restriction policies (UK Annex | TikTok, 2025). The final dataset comprised 741 videos with complete multimodal feature extraction. Appendix 1 provides a detailed breakdown of transcription sources and video counts.

Data Sources

Three sources provided the data: the TikTok Research API, a third-party analytics service (Analisa.io), and manual verification procedures. The first dataset was acquired in March 2023 from Analisa.io, a commercial analytics service (Lachief, 2023); access to several creator profiles was later lost following its closure. The second dataset was collected in April 2024 via the official TikTok Research API, accessed using R packages

including httr and jsonlite (Pascual, 2020). Data were filtered by publication date (2020 to 2022), geographic region (United Kingdom and United States), and key content metrics detailed in Appendix 2. Manual verification occurred between April and July 2023, identifying missing videos between website metadata and API-sourced data.

Neither source provided native video files. Video files were therefore retrieved from their publicly available TikTok URLs using third-party services. Audio tracks were subsequently extracted from each file, saved in MP4 format, and converted to WAV prior to acoustic analysis.

The University of Essex granted ethical approval (Reference: ETH2526-0311). All data were collected from publicly available TikTok content voluntarily posted by creators and accessed via the TikTok Research API. Individual creator consent was not required under applicable institutional guidelines.

3.2 Variables and Measures

This section describes all variables and measurement procedures, organised by type. Variables shared across both research questions are introduced here and not repeated thereafter (Variable descriptors can be found in Table 2).

Dependent Variables

For RQ1, a comprehensive set of multimodal production features served as the dependent variables, described in the subsection 3.3. For RQ2, three engagement metrics served as outcome measures: like count, comment count, and share count. These metrics reflect qualitatively distinct forms of audience response. Likes indicate passive approval; comments signal active discursive participation; and shares represent content amplification through network diffusion (Habibi & Salim, 2021). Log-transformed view count was included in all RQ2 models as a control variable to account for differences in baseline

Independent Variables

Independent variables spanned five dimensions: creator-level features, visual features, audio features, linguistic features, and platform features. Creator-level features comprised creator gender (male/female) and nationality (UK/US/other). These were included to account for demographic characteristics that may systematically influence content style and audience reception (Bhandari & Bimo, 2022).

Visual features comprised video duration (in seconds), brightness, contrast, rule-of-thirds alignment score (compositional balance metric), presence of visual filters, and use of green screen technology. Brightness and contrast were derived from pixel luminance values on a normalised scale of zero to one (Bilozerskyi et al., 2023; Schwenzow et al., 2021; Watier, 2023). The rule-of-thirds score measures how closely salient image elements align with the intersections of a three-by-three grid (Jin et al., 2022). Higher scores indicate more deliberate compositional structure, whilst scores approaching zero suggest centred framing or images with few distinct edges (Zhang et al., 2023).

Audio features included RMS (root mean square) amplitude, zero-crossing rate, mean pitch, pitch variability (standard deviation of pitch), speech rate, background music, sound effects, and accent. RMS amplitude measures overall loudness, with larger values indicating louder audio energy (Er, 2020; Hove et al., 2019). Zero-crossing rate measures the frequency with which an audio signal transitions from positive to negative values per unit of time (Er, 2020; Rehman et al., 2021). Higher values correspond to noisier or brighter sound textures (ibid). Mean pitch and pitch variability were computed from voiced segments, converted to semitones using the formula:

$$ST = 12 \times \log_2\left(\frac{Hz}{50}\right),$$

following established prosody research practice (Shi et al., 2022). Speech rate was calculated as syllables per second.

Linguistic features comprised type-token ratio (TTR), BNC word frequency, transcript sentiment, title sentiment, video description word count, first-person pronoun ratio, and second-person pronoun ratio. TTR assessed lexical diversity by dividing the number of unique word types by total tokens. Higher values indicate greater vocabulary variation. BNC word frequency was calculated by comparing transcription vocabulary against the British National Corpus frequency list, yielding a measure of how common or specialised the language was list (Leech et al., 2014). Sentiment was analysed using the **syuzhet** package in R (Jockers, 2015), yielding continuous scores in which positive values indicate positive valence. Separate scores were computed for video transcriptions and video titles. First-person pronoun ratio (I, me, my, mine, myself) and second-person pronoun ratio (you, your, yours, yourself, yourselves) were each standardised by total transcript

word count (Cheng & Li, 2024; Chung et al., 2023). Platform features comprised use of TikTok's duet and stitch functions, which enable collaborative or responsive content creation, and use of visual filters.

Role-Play Variables

Role-play was first coded as a binary indicator denoting whether a creator performed a character distinct from their everyday persona within a given video. A derived variable, `role_play_type`, distinguished three levels: `No_RolePlay` (baseline), `SameGender` (same-gender character portrayal), and `CrossGender` (cross-gender character portrayal). Coding distinguished same-gender from cross-gender role-play to reflect their distinct performative demands and potential for differential audience responses.

Coding strategy. Role-play was identified through direct visual observation of each video. A video was coded as role-play when the creator explicitly adopted a character persona distinct from their everyday on-screen identity, as evidenced through at least two of the following observable cues: a change in voice quality or accent, a change in costume or accessories, a distinct gestural or bodily manner, or an explicitly fictional narrative framing (for instance, a scripted scenario with dialogue between characters). Videos in which creators spoke directly to camera in their own persona, without adopting a distinct character, were coded as `No_RolePlay`. Where the role-play classification was ambiguous from the video alone, the video description and transcript were consulted to confirm whether a character persona was being enacted.

Classification as `SameGender` or `CrossGender` was based on whether the character's gender presentation matched or differed from the creator's own gender. `CrossGender` coding required explicit embodiment of gendered markers through at least two channels simultaneously, specifically voice modulation consistent with a different gender, gender-coded costume or accessories (such as a wig, skirt, or stereotypically feminine or masculine clothing), or gestural and postural cues associated with a different gender. Videos in which a single channel of gender-coded performance was observable, or in which gender cues were ambiguous, were classified as `SameGender`. This threshold ensured that `CrossGender` coding reflected deliberate and multimodally constructed gender-crossing rather than isolated stylistic choices. All role-play coding was conducted by the researcher through direct observation, as the features in question are explicitly signalled

within the video content itself and are identifiable through perceptual observation of the video, description, and transcript without requiring inferential judgement.

Distinguishing same-gender from cross-gender role-play is theoretically motivated. Cross-gender role-play introduces a distinct layer of performance, requiring explicit embodiment of gendered cues through voice, costume, and gesture. These additional demands may generate qualitatively different multimodal signatures and may elicit different audience responses through novelty and expectation violation (Wu et al., 2025). Collapsing the two into a single role-play category would obscure these distinctions.

Role-play-specific production elements captured features integral to character construction. These comprised costume design (clothing and accessories), makeup and hairstyling (prosthetic makeup and different hairstyles), and visual design (green screen use and setting changes). These features served as dependent variables in RQ1 and as additional predictors in the full RQ2 model specification.

Measurement Procedure

The `imager` and `magick` packages in R (Barthelme, 2015; Ooms, 2016) extracted continuous visual features frame by frame, with mean values computed across each video (Figure 1). The `tuner` and `seewave` packages (Ligges et al., 2004; Sueur et al., 2006) yielded audio features. Linguistic features were computed from video transcriptions. Video duration was returned directly by the API.

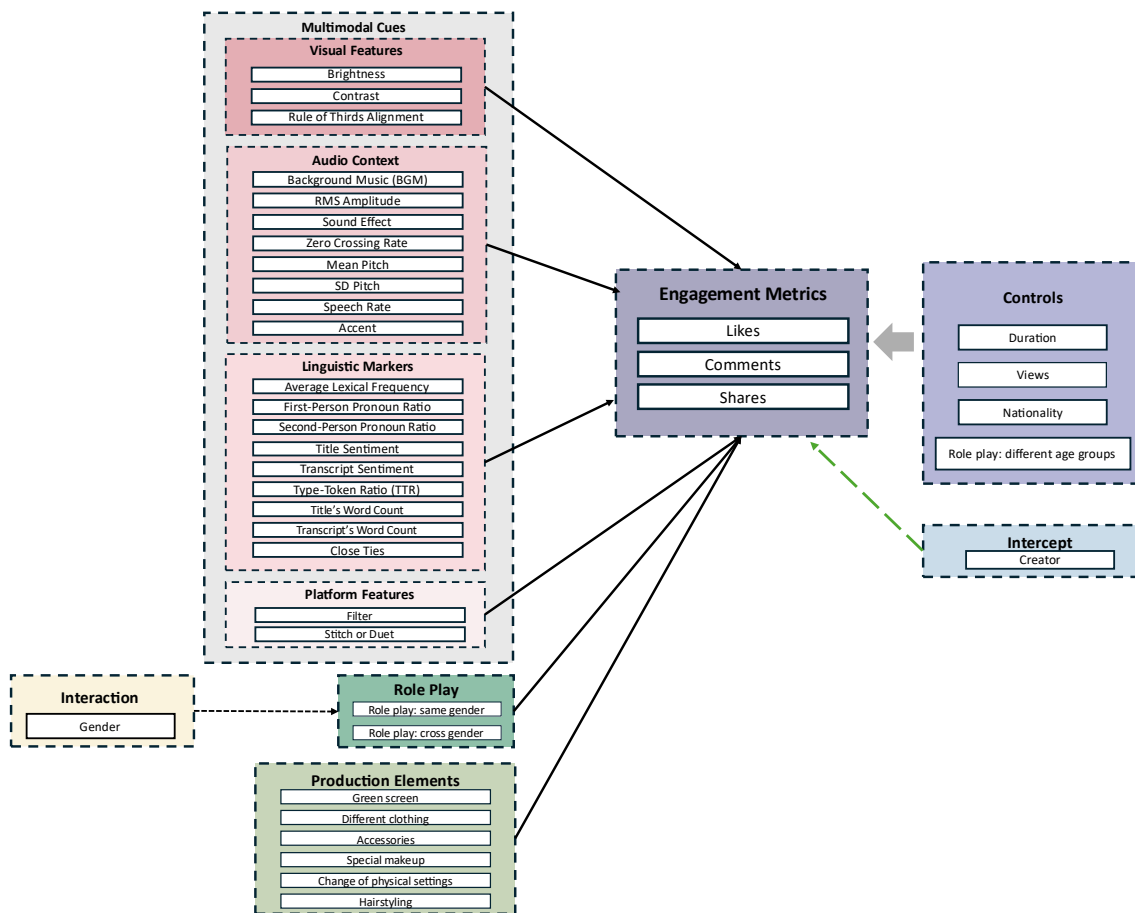
Binary production features were coded through direct observation of each video. These comprised filters, green screen, background music, sound effects, accent, costume and accessories, hairstyling, makeup, physical settings, stitch and duet use, and role-play variables. Each feature is perceptually grounded and operationally explicit. Accessories such as wigs are visually identifiable in the video frame. Stitch and duet use is apparent from the split-screen format or watermark overlay. Green screen use produces a visually distinctive background effect. Where additional context was required to confirm a coding decision, video descriptions and transcripts were consulted.

The coding criteria for binary features are defined by the verifiable presence or absence of specific, discrete visual or platform elements. Several features are effectively platform-coded. Stitch and duet use is identifiable from the split-screen format or watermark overlay. Green screen use produces a visually distinctive background

effect. Others are directly legible from video subtitles or descriptions without requiring perceptual inference. This operational specificity substantially constrains interpretive latitude. Coding decisions of this kind differ from constructs requiring subjective evaluation, where interpretive variation between observers is unavoidable. Automated extraction cannot reliably identify platform-specific affordances such as filter use or stitch and duet functions, which motivated the manual approach.

Creator gender and nationality were included as covariates in all models. All models additionally included creator-specific random intercepts to address the nested structure of videos within creators.

Figure 22 Research Framework of Multimodal Models



Variable Transformation

Transformations were applied to ensure variables were on comparable scales and that model assumptions were met. Preliminary examination revealed substantial positive skewness in engagement metrics and several predictor variables, necessitating logarithmic transformations to approximate normal distributions and satisfy regression assumptions. Engagement outcomes (likes, views, comments, and shares) were log-transformed using $\log(x + 1)$ to reduce skewness and stabilise variance (Gelman & Hill, 2006). Duration and count variables

were log-transformed and then standardised. Proportional measures were also log-transformed where skewness was pronounced, then standardised.

For prosodic cues, speech rate was log-transformed, and raw pitch values were converted from Hertz into semitones relative to a 50 Hz reference using the formula (Shi, Gu, and Vigliocco, 2022):

$$ST = 12 \times \log_2\left(\frac{\text{Hz}}{50}\right),$$

Both semitone measures were subsequently standardised (z-scored) prior to analysis. Negatively skewed variables, including average lexical frequency and brightness, were squared before standardisation (Tabachnick & Fidell, 2019). Sentiment scores were centred and scaled only, as they are bounded around zero and unsuitable for log transformation (Smithson & Verkuilen, 2006; Tabachnick & Fidell, 2019). Categorical variables, including nationality, gender, background music, filters, sound effects, stitch or duet use, and close ties, were converted to factors. This transformation strategy ensured that all predictors contributed appropriately to the statistical models whilst maintaining interpretability of effect sizes.

3.3 Data Analysis

Data Diagnostics

Bivariate diagnostics are reported in full in the Appendix 3 and 4. In brief, the three engagement metrics were strongly and significantly intercorrelated: likes and shares ($r = 0.76$), likes and comments ($r = 0.70$), and shares and comments ($r = 0.77$), all $p < .001$. This confirms that popular content tends to elicit multiple forms of audience response simultaneously.

Several predictor pairs exhibited notable multicollinearity. The most pronounced instances were speech rate and type-token ratio ($r = 0.55$), duration and type-token ratio ($r = 0.54$), and mean pitch and zero-crossing rate ($r = 0.67$). These were addressed through variable transformation and model specification prior to analysis. Variance inflation factors confirmed the absence of problematic multicollinearity in final models, with all VIF values below 2.5 following removal of collinear variables (Mfondoum et al., 2023). Transcript word count was removed owing to collinearity with duration.

Analytic Strategy

The analytical approach employed distinct methods for each research question. For RQ1, LMMs were fitted for continuous production element variables and GLMMs with logistic link functions for binary variables. All RQ1 models included creator-specific random intercepts to account for the nested structure of videos within creators. Each model tested main effects of `role_play_type` (No_RolePlay, SameGender, CrossGender) and creator gender, alongside their interaction term.

RQ1: Production Element Differences

Interaction terms were included because production choices may reflect not only role-play demands but also gender-specific affordances in character construction. Male and female creators may draw upon divergent vocal ranges, costume conventions, or gestural repertoires when constructing characters. The model specification took the form:

$$Y = \text{role_play_type} \times \text{gender} + (1|\text{creator})$$

where Y represents each production element in turn. Interaction effects were interpreted through predicted values and simple slopes, comparing role-play types against the no-role-play baseline within each gender group. False Discovery Rate correction was applied using the Benjamini-Hochberg method to control for multiple comparisons across production elements.

RQ2: Engagement Prediction

The relationship between continuous predictors and engagement outcomes cannot be assumed to be linear a priori. Production variables such as video duration, pitch, and speech rate may exhibit threshold or saturation effects. Generalised Additive Mixed Models (GAMMs) were therefore employed with creator-specific random intercepts to model log-transformed engagement outcomes, using thin-plate regression splines with REML estimation and basis dimension $k = 5$ (Wood, 2017).

Model comparison proceeded in four stages, with results for each engagement metric reported correspondingly. The **baseline model** included `role_play_type` alongside all technical production elements (visual, audio, linguistic features) and control variables (nationality, gender, view count), with **role-play effects** and **production element effects** reported from this specification. An **interaction model** added `role_play_type` $\times$ gender terms to test whether engagement effects differ by creator gender. The **full model** incorporated role-

play-specific production elements (green screen, clothing, accessories, hairstyling, makeup, physical settings, accent). Model fit was evaluated using AIC, with lower values indicating superior fit.

The inclusion of gender interactions was theoretically motivated. Gender stereotypes establish culturally shared expectations about how men and women perform, and departures from those expectations are processed differently depending on the gender of the person departing (Eagly & Karau, 2002; Morgenroth & Ryan, 2021). Butler's (1990) account of gender performativity further suggests that the cultural weight of performative acts differs by performer gender: the same act of gender-norm transgression carries different social meanings when enacted by a male versus a female creator. Empirical work corroborates this: Wu et al. (2025) demonstrate that gender-incongruent content on microblogs generates higher engagement, and that the magnitude of this effect differs between male and female creators. Within a comedic context specifically, the social licence for gender transgression is amplified by the performative frame, which signals to audiences that the norms being cited may also be subverted. Interaction terms were therefore included to enable separate contrasts for same-gender and cross-gender role-play against the no-role-play baseline within each gender group. All analyses were conducted in R version 4.3.2 using lme4, gamm4, mgcv, and performance packages (Lüdtke et al., 2019; Wood & Scheipl, 2009; Wood, 2000).

4. Results

4.1 Descriptive Statistics

Table 2 presents the descriptive statistics for all variables included in the analysis. Bivariate diagnostics (Appendix 3 and 4) revealed that the three-engagement metrics were strongly and significantly intercorrelated (likes–shares: $r = 0.76$; likes–comments: $r = 0.70$; shares–comments: $r = 0.77$; all $p < .001$), confirming that popular content tends to elicit multiple forms of audience response simultaneously. Several predictor pairs exhibited notable multicollinearity, most notably speech rate and type-token ratio ($r = -0.55$), duration and type-token ratio ($r = -0.54$), and mean pitch and zero-crossing rate ($r = -0.67$).

Table 11 Table 2 Descriptive statistics of variables

Variables	N	Mean	Std	Min	Max	Skewness	Category	Descriptions
Comments	788	90676.02	830359.08	0.00	12500000	11.13	Engagement, DV	The count of comments of a TikTok video
Shares	788	17184.84	59246.72	4.00	1300000	14.20	Engagement, DV	The count of shares of a TikTok video
Likes	788	601016.29	1008094.24	180	8500000	3.65	Engagement, DV	The count of likes of a TikTok video
Views	788	5602318.39	12158358.13	4734	163900000	6.48	Engagement, IV	The count of views of a TikTok video
nationality	788	1.51	0.58	1.00	3.00	0.65	Creator-Level Features	1=American, 2=British, 3=Chinese-Irish
gender	788	1.67	0.47	1.00	2.00	-0.71	Creator-Level Features	1=female, 2=male
duration	788	49.38	33.85	4.00	351	3.40	Video	how long the video is in seconds
brightness	741	0.48	0.12	0.00	0.86	-0.81	Video	the overall lightness or luminance of the frames
contrast	741	0.23	0.04	0.11	0.35	0.06	Video	the difference in luminance (light intensity) between the darkest and brightest parts of the frames.
rule.of.thirds_alignment	741	0.00	0.000383	0.000073	0.005561	4.37	Video	Frames aligned with the rule of thirds are often perceived as more professional and visually engaging. A higher rule-of-thirds score indicates more aesthetically pleasing composition, whilst scores closer to zero suggest centred subjects or images with few distinct edges.
green_screen	788	1.06	0.23	1.00	2.00	3.81	Video	binary: whether the video uses green screen or not (No=1, Yes=2)
bgm	788	1.02	0.13	1.00	2.00	7.58	Audio	binary: whether the video uses background music or not (No=1, Yes=2)
speech_rate_sylls	788	2.29	1.09	0.07	6.57	0.07	Audio	syllables per second (syll/s)
sound_effect	788	1.02	0.13	1.00	2.00	7.58	Audio	binary: whether the video uses sound effects or not (No=1, Yes=2)
rms_amplitude	741	3725.08	2016.44	136.92	13164.26	1.44	Audio	Measuring loudness of speech or music in raw PCM amplitude units
zero_crossing_rate	741	0.06	0.02	0.02	0.16	0.60	Audio	the frequency with which an audio signal transitions
mean_pitch	741	7.20	7.03	1.21	14.52	0.42	Audio	The average pitch of the given audio file in Hz
sd_pitch	741	4.88	4.94	1.23	7.05	-0.74	Audio	Standard deviation of pitch values of the given audio file: a small SD means the voice is relatively stable in pitch in Hz
accent	788	1.04	0.19	1.00	2.00	4.91	Audio	binary: whether the creators use a different accent from theirs (No=1, Yes=2)
ttr	788	0.66	0.16	0.21	1.00	0.50	Linguistic markers	Type-token ratio: A higher TTR means the text uses a wider variety of words relative to its length (lexically rich)
avg_lexical_frequency	788	6.49	0.29	0.00	7.41	-13.84	Linguistic markers	The average number of times a given word in one transcript appears in BNC frequency list
transcript_senti	788	0.90	2.71	-8.30	14.1	0.50	Linguistic markers	sentiment scores of a transcript: positive value indicates positive sentiment and vice versa
title_senti	788	0.15	0.75	-2.00	4.65	1.06	Linguistic markers	sentiment scores of a video description: positive value indicates positive sentiment and vice versa
word_count_title	788	9.76	6.96	1.00	85	3.37	Linguistic markers	word count of a video description
word_count_transcript	788	112.99	83.33	3.00	656	1.28	Linguistic markers	word count of a transcription
close_ties	788	1.25	0.43	1.00	2.00	1.15	Linguistic markers	binary: Yes-videos concerning family, friendship and relationship (Chung et al., 2023); No-the rest of the topics (No=1, Yes=2)
filter	788	0.09	0.05	0.00	0.38	3.62	Platform features	binary: whether the video uses filter or not (No=1, Yes=2)
stitch_duet	788	1.06	0.24	1.00	2.00	4.15	Platform features	binary: whether the video uses stitch/duet editing features from TikTok or not (No=1, Yes=2)
clothing	788	1.05	0.22	1.00	2.00	1.01	Production elements	binary: whether different clothes are applied for different characters or not (No=1, Yes=2)
accessories	788	1.27	0.45	1.00	2.00	0.18	Production elements	binary: presence of extra props (No=1, Yes=2)
hairstyling	788	1.45	0.5	1.00	2.00	2.80	Production elements	binary: whether different hairstyles are applied for different characters or not (No=1, Yes=2)
makeup	788	1.09	0.29	1.00	2.00	7.03	Production elements	binary: presence of prosthetic makeup (No=1, Yes=2)
physical_settings	788	1.02	0.14	1.00	2.00	1.00	Production elements	binary: whether different settings are applied for different scenes/characters or not (No=1, Yes=2)
role_play	788	1.28	0.45	1.00	2.00	-0.30	Role play characteristics	binary: presence/absence of creator playing a role (No=1, Yes=2)
diff_gender	788	1.57	0.49	1.00	2.00	3.16	Role play characteristics	binary: presence/absence of cross-gender performance (No=1, Yes=2)

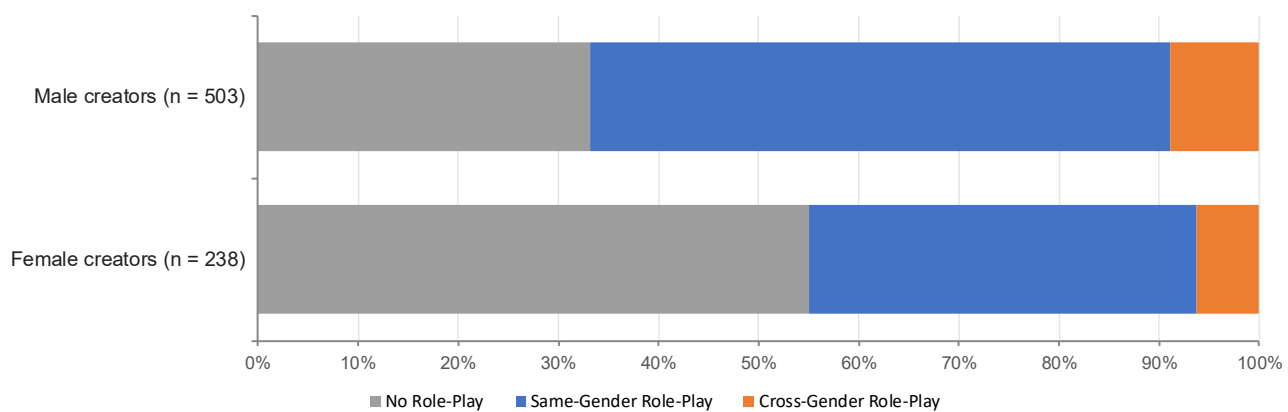
diff_age	788	1.08	0.27	1.00	2.00	2.66	Role play characteristics	binary: presence/absence of cross-age performance (No=1, Yes=2)
role_play_type	741	1.10	0.3	1.00	2.00	0.33	Role play characteristics	1=No_RolePlay, 2= SameGender, 3=CrossGender

Note: The dataset comprised 788 TikTok videos, with slight variations in sample size for certain audio-visual features (n = 741) due to technical constraints in feature extraction

4.2 Production Element Differences by Role-Play Type (RQ1)

Table 3 presents gender-differentiated effects for production elements showing significant interactions ($p < .10$). For continuous variables, coefficients represent the effect of role-play type compared to non-role-play baseline. For binary variables, odds ratios are presented. Interaction terms indicate how the effect differs for male creators relative to female creators (reference group). The sample comprised 741 videos: 298 non-role-play, 383 same-gender role-play, and 60 cross-gender role-play. Figure 2 presents the distribution by creator gender. All odds ratios compare non-role-play content as the baseline. Interaction estimates are relative to female creators unless stated otherwise. All odds ratios compare against non-role-play content as the baseline. Interaction estimates are relative to female creators unless stated otherwise.

Figure 23 Distribution of comedic TikTok videos by role-play type and creator gender (N = 741).



Role-Play Type	Female creators	Male creators	Total
No Role-Play	131 (55.0%)	167 (33.2%)	298 (40.2%)
Same-Gender Role-Play	92 (38.7%)	291 (57.9%)	383 (51.7%)
Cross-Gender Role-Play	15 (6.3%)	45 (8.9%)	60 (8.1%)
Total	238	503	741

Visual Elements

Role-play videos showed marked increases in costume and styling elements. Accessories demonstrated the strongest effect. Same-gender role-play was 90 times more likely to feature accessories than non-role-play (OR = 90.46, $p < .001$, FDR $p < .001$). Cross-gender role-play showed similar odds (OR = 91.54, $p < .001$, FDR $p < .001$).

Clothing differentiation followed comparable patterns. Same-gender videos were 4.2 times more likely to feature distinct costumes (OR = 4.21, $p < .001$, FDR $p = .005$). Cross-gender portrayals showed substantially

higher odds ($OR = 36.33, p < .001, FDR p < .001$). Physical settings differentiated content types. Same-gender role-play was 3.1 times more likely to feature varied locations ($OR = 3.05, p = .010, FDR p = .017$).

Technical visual features showed substantial elevation. Filters were 22.70 times more prevalent in same-gender role-play ($OR = 22.70, p < .001, FDR p < .001$). Cross-gender videos showed 8.32 times higher filter prevalence ($OR = 8.32, p < .001, FDR p < .001$). Green screen adoption demonstrated stronger patterns. Same-gender role-play was 20.18 times more likely to employ this technology ($OR = 20.18, p = .006, FDR p = .013$). Cross-gender videos showed even stronger effects ($OR = 28.54, p < .001, FDR p < .001$).

Hairstyling and makeup associations did not survive FDR correction. Same-gender hairstyling showed $p = .745$ ($FDR p = .894$) and cross-gender $p = .694$ ($FDR p = .925$). For makeup, same-gender results were $p = .592$ ($FDR p = .789$) and cross-gender $p = .940$ ($FDR p > .999$). The sparse occurrence of prosthetic makeup (15 videos) and specialised hairstyling (under 13% of same-gender role-play) produced unstable estimates. Brightness, contrast, and rule of thirds showed no significant differences. Visual composition remained relatively consistent across role-play types.

Gender interactions emerged primarily in technical features. Male creators employing same-gender role-play used filters substantially more than female creators ($OR = 27.41, p = .011$). Green screen usage approached significance ($OR = 15.73, p = .086$), with elevated adoption among male same-gender creators. Accessories, clothing, hairstyling, and physical settings showed minimal interactions (all $p > .10$). These choices reflect role-play demands rather than gender-specific preferences.

Audio Elements

Role-play videos exhibited distinctive audio characteristics. RMS amplitude showed notable negative associations with role-play. Same-gender role-play demonstrated substantially lower loudness than baseline content ($\beta = -906.69, SE = 249.54, p < .001, FDR p = .003$). Cross-gender role-play showed even stronger effects ($\beta = -1178.30, SE = 286.07, p < .001, FDR p < .001$). Zero crossing rate likewise differentiated role-play content. Same-gender videos showed reduced audio brightness relative to baseline ($p < .001$).

Background music was 11.49 times more prevalent in same-gender role-play ($OR = 11.49, p < .001, FDR p < .001$). Same-gender videos were 12.67 times more likely to employ stitch or duet functions ($OR = 12.67, p$

= .022, FDR p = .033). Sound effects, mean pitch, pitch variability, and speech rate showed no significant effects after FDR correction.

Gender interactions revealed differentiated audio strategies between creators. Male creators were substantially more likely to incorporate background music in same-gender role-play (OR = 14.37, p = .003). Zero crossing rate showed significant interactions for both same-gender (p < .001) and cross-gender role-play (p = .006). The audio brightness reduction associated with role-play was significantly attenuated amongst male creators. Mean pitch modulation also varied by gender (p = .037 for same-gender). These patterns likely reflect gender-specific vocal affordances.

Table 12 Gender by Role-Play Type Interactions for Production Elements

Continuous Production Elements						
Production Element	Role-Play Type	Female Creators (Reference)		Male Creators		Gender Interaction
		β (SE)		β (SE)	β (SE)	p
<i>Linguistic/Textual</i>						
Duration (seconds)	Same-gender	29.43 (4.65)	***	3.41 (5.70)	26.02 (7.36)	< .001***
	Cross-gender	11.21 (5.38)	*	-7.22 (9.59)	18.43 (10.99)	0.094†
Word Count (Title)	Same-gender	3.87 (1.21)	**	-0.02 (1.45)	-3.89 (1.55)	0.013*
	Cross-gender	7.39 (1.56)	***	6.10 (2.25)	-1.29 (2.31)	0.578
Title Sentiment	Same-gender	0.53 (0.12)	***	0.04 (0.14)	-0.48 (0.17)	0.004**
	Cross-gender	0.40 (0.18)	*	0.47 (0.24)	0.07 (0.25)	0.780
First-Person Pronoun Ratio	Same-gender	-0.010 (0.003)	**	0.004 (0.004)	-0.015 (0.005)	0.006**
	Cross-gender	-0.007 (0.005)		0.006 (0.009)	-0.012 (0.010)	0.233
<i>Audio</i>						
Zero Crossing Rate	Same-gender	-0.013 (0.003)	***	-0.032 (0.005)	-0.018 (0.004)	< .001***
	Cross-gender	-0.012 (0.005)	*	-0.030 (0.008)	-0.018 (0.007)	0.006**
Mean Pitch (semitones)	Same-gender	0.20 (0.29)		1.03 (0.35)	0.83 (0.40)	0.037*
	Cross-gender	0.78 (0.44)		1.63 (0.59)	0.85 (0.59)	0.151
SD Pitch (semitones)	Same-gender	0.18 (0.13)		0.52 (0.18)	0.34 (0.17)	0.049*
	Cross-gender	0.24 (0.19)		0.36 (0.27)	0.13 (0.26)	0.629
RMS Amplitude	Same-gender	-906.69 (249.54)	***	-407.27 (308.54)	499.42 (396.82)	0.209
	Cross-gender	-1178.30 (286.07)	***	-503.23 (513.63)	675.07 (587.92)	0.251
<i>Visual</i>						
Contrast	Same-gender	0.009 (0.006)		0.007 (0.007)	-0.002 (0.008)	0.775
	Cross-gender	0.010 (0.006)		-0.014 (0.010)	0.023 (0.012)	0.047*
Brightness	Same-gender	0.013 (0.019)		-0.005 (0.023)	-0.018 (0.025)	0.468
	Cross-gender	0.065 (0.035)		0.015 (0.048)	-0.049 (0.040)	0.223
Binary Production Elements						
		OR	(95% CI)	OR	(95% CI)	p
<i>Visual/Technical</i>						
Filter	Same-gender	22.7	(5.49–93.90)***	622.15	(79.87–4846.60)***	0.011*
	Cross-gender	8.32	(1.95–35.43)**	368.47	(4.74–28,648.80)**	0.773
Green Screen	Same-gender	20.18	(2.43–167.80)**	317.31	(21.35–4716.40)***	0.086†
	Cross-gender	28.54	(4.47–182.30)***	24,001,800	(CI not estimable)***	0.743
<i>Audio</i>						
Background Music	Same-gender	11.49	(2.82–46.78)***	165.02	(28.91–942.00)***	0.003**
	Cross-gender	7.12	(0.98–51.58)†	42.14	(2.14–828.20)*	0.331
<i>Costume/Styling</i>						
Accessories	Same-gender	90.46	(20.09–407.40)***	27.76	(3.23–238.70)**	0.329
	Cross-gender	91.54	(12.44–673.60)***	13.55	(0.85–216.00)	0.166
Clothing	Same-gender	4.21	(1.66–10.70)**	4.97	(1.48–16.70)**	0.845
	Cross-gender	36.33	(7.47–176.60)***	63.82	(6.13–664.80)***	0.646
Hairstyling	Same-gender	3.66	(0.55–24.50)	340.89	(3.03–38,353.10)**	0.846
	Cross-gender	21.8	(2.63–180.70)**	20,946,000	(CI not estimable)***	0.827
Physical Settings	Same-gender	3.05	(1.31–7.12)**	1.82	(0.57–5.81)	0.510
	Cross-gender	0.3	(0.03–2.66)	0.17	(0.01–2.69)	0.751
<i>Narrative</i>						
Close Ties	Same-gender	7.35	(2.86–18.88)***	4.21	(1.33–13.36)*	0.426
	Cross-gender	26.6	(6.43–110.00)***	36.32	(3.99–330.20)***	0.734

Note:

Female creators serve as the reference category. Interaction terms indicate the multiplicative change in effect for male creators relative to female creators.

For continuous variables: Positive interaction coefficients indicate that male creators show a larger positive effect (or smaller negative effect) than female creators. Negative coefficients indicate that male creators show a smaller positive effect (or larger negative effect).

For binary variables: Interaction ORs > 1 indicate that male creators show stronger associations than female creators. ORs < 1 indicate weaker associations for male creators.

Significance levels: † $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

CI = Confidence Interval; OR = Odds Ratio; SE = Standard Error

Some OR estimates for binary variables exceed typical ranges due to sparse data in specific cells, particularly for green screen and hairstyling in cross-gender role-play. Where noted as "CI not estimable," confidence intervals could not be reliably calculated due to extremely sparse data (zero or near-zero cell counts), resulting in numerical instability. These estimates should be interpreted with caution and are primarily reported for completeness.

Linguistic and Textual Elements

Title word count differentiated role-play from non-role-play content. Same-gender videos featured longer titles ($\beta = 3.87$, $SE = 1.21$, $p = .001$). Cross-gender videos showed even greater elevation ($\beta = 7.39$, $SE = 1.56$, $p < .001$). These effects remained notable under stricter correction thresholds. Transcript word count, mean length of utterance, type-token ratio, and sentiment measures showed no significant main effects. Pronoun ratios likewise showed no significant associations after multiple comparison adjustment.

Gender interactions emerged in linguistic choices across role-play types. Male creators produced substantially longer same-gender role-play videos than female creators ($\beta = 26.02$, $SE = 7.36$, $p < .001$). Cross-gender role-play showed a similar though weaker pattern ($\beta = 18.43$, $SE = 10.99$, $p = .094$). Title word count demonstrated a negative interaction for same-gender role-play ($\beta = -3.89$, $SE = 1.55$, $p = .013$). Title sentiment revealed gender differentiation ($\beta = -0.48$, $SE = 0.17$, $p = .004$). Male creators employed different affective framings than female creators. First-person pronoun usage showed significant interactions ($p = .006$). Male creators reduced self-reference in role-play more than female creators.

Narrative Content Elements

Close ties content was substantially more prevalent in role-play videos. Same-gender performances were 7.3 times more likely to address relationship themes ($OR = 7.35$, $FDR\ p < .001$). Cross-gender role-play showed even stronger associations ($OR = 26.60$, $FDR\ p < .001$). Age differentiation between characters showed no significant associations with role-play type. Gender interactions for close ties were minimal ($p = .426$). This narrative choice operates independently of creator gender.

Engagement Metrics

Engagement metrics did not differ significantly by role-play type in preliminary analyses. Likes showed no reliable differences ($p = .156$). Comments likewise showed no significant differences ($p = .201$). Shares also showed no reliable differences ($p = .087$). Role-play presence alone did not determine audience engagement. Interactive effects with other production elements appear necessary for observable engagement differences.

4.3 Production Elements and Engagement Outcomes (RQ2)

4.3.1 Like Count

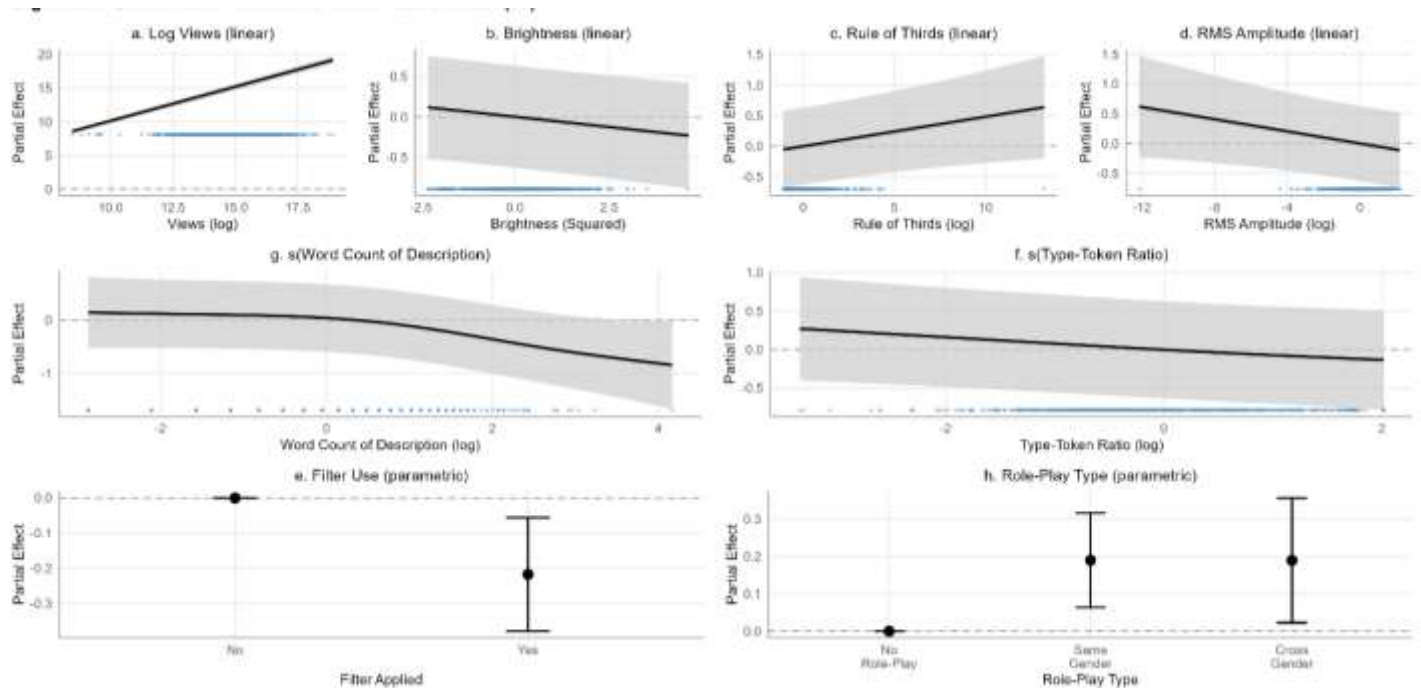
The baseline model (Figure 2; Model 1a, Table 4) demonstrated strong predictive performance (adjusted $R^2 = 0.734$, AIC = 1290.99). Views showed strong positive association with likes ($\beta = 1.014$, SE = 0.016, $p < .001$). Same-gender role-play significantly predicted elevated likes ($\beta = 0.190$, SE = 0.064, $p = .003$). Cross-gender role-play likewise increased likes ($\beta = 0.189$, SE = 0.085, $p = .027$). Each represented approximately 21% more likes net of production elements.

Several production features independently predicted likes (see Figure 2). Description length was the strongest predictor (edf = 2.681, $F = 13.214$, $p < .001$), with diminishing returns at higher word counts. Greater lexical diversity predicted more likes (edf = 1.261, $F = 6.026$, $p = .028$). Higher brightness increased likes (edf = 1.000, $F = 7.263$, $p = .007$). Stronger rule-of-thirds composition also increased likes (edf = 1.000, $F = 4.855$, $p = .028$). Lower loudness was another positive predictor (edf = 1.000, $F = 4.481$, $p = .035$). Filter use predicted lower engagement ($\beta = -0.218$, SE = 0.082, $p = .009$).

An interaction model introducing role-play type by gender yielded marginally worse fit (adjusted $R^2 = 0.730$, AIC = 1306.338). The same-gender role-play by male creator interaction suggested modest benefits for male creators ($\beta = 0.211$, SE = 0.129, $p = .104$). This effect did not reach significance. Model comparison favoured the simpler additive specification based on AIC.

Adding role-play-specific production elements degraded fit (adjusted $R^2 = 0.736$, AIC = 1333.39). Role-play effects remained robust (same-gender: $\beta = 0.200$, SE = 0.073, $p = .006$; cross-gender: $\beta = 0.223$, SE = 0.098, $p = .023$). These costume and visual features contributed little beyond role-play type itself.

Figure 24 Significant Smooth and Parametric Terms in Likes Model (1a)



"Note: Brightness (Squared): squared brightness value.

"Rule of Thirds (log): log-transformed rule-of-thirds alignment score.

"RMS Amplitude (log): log-transformed Root Mean Square amplitude, indicating loudness.

"Type-Token Ratio (log): log-transformed type-token ratio, indicating lexical diversity.

"Word Count of Description (log): log-transformed word count of the video description.

"Filter Use and Role-Play Type: parametric coefficients relative to reference category.

4.3.2 Comment Count

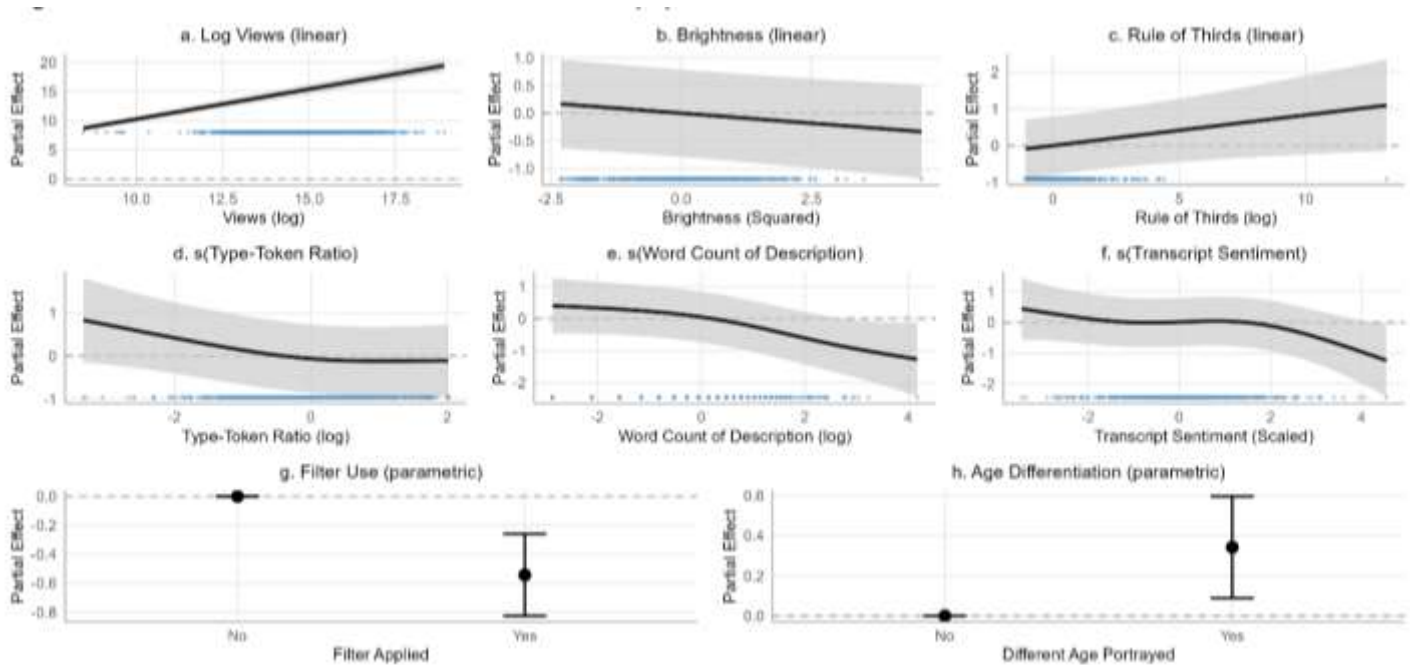
The baseline model (Figure 3; Model 2a, Table 4) showed moderate predictive performance (adjusted $R^2 = 0.618$, AIC = 2091.766). Views demonstrated strong positive effect ($\beta = 1.027$, SE = 0.028, $p < .001$). Yet same-gender role-play did not predict comments ($\beta = -0.017$, SE = 0.112, $p = .877$). Cross-gender role-play also showed no significant effect ($\beta = 0.201$, SE = 0.149, $p = .180$). Role-play type does not independently drive commenting behaviour when production elements are controlled.

Several production features shaped commenting. Videos with longer descriptions attracted substantially more comments (edf = 2.317, $F = 17.567$, $p < .001$). Greater lexical richness predicted discussion (edf = 2.254, $F = 5.885$, $p = .002$). Moderate affective valence in speech stimulated comments (edf = 3.072, $F = 3.469$, $p = .015$). Moderately valenced content drew more comments than neutral or extreme material. Brighter videos increased comments (edf = 1.000, $F = 4.906$, $p = .027$). Stronger compositional balance also increased comments (edf = 1.000, $F = 4.725$, $p = .030$).

Filter use predicted substantially fewer comments ($\beta = -0.542$, $SE = 0.144$, $p < .001$). Cross-age character portrayals increased discussion ($\beta = 0.342$, $SE = 0.129$, $p = .008$). Intergenerational scenarios may invoke relatable family dynamics or mentorship experiences.

An interaction model introducing role-play type by gender produced marginal improvement (adjusted $R^2 = 0.615$, $AIC = 2094.538$). The same-gender role-play by male creator interaction approached significance ($\beta = 0.394$, $SE = 0.225$, $p = .081$). The full model degraded fit (adjusted $R^2 = 0.617$, $AIC = 2115.649$), with role-play type remaining non-significant throughout all specifications.

Figure 25 Significant Smooth and Parametric Terms in Comments Model (2a)



Note: *Brightness (Squared)*: squared brightness value.

Rule of Thirds (log): log-transformed rule-of-thirds alignment score.

RMS Amplitude (log): log-transformed Root Mean Square amplitude, indicating loudness.

Type-Token Ratio (log): log-transformed type-token ratio, indicating lexical diversity.

Word Count of Description (log): log-transformed word count of the video description.

Filter Use and Role-Play Type: parametric coefficients relative to reference category.

4.3.3 Share Count

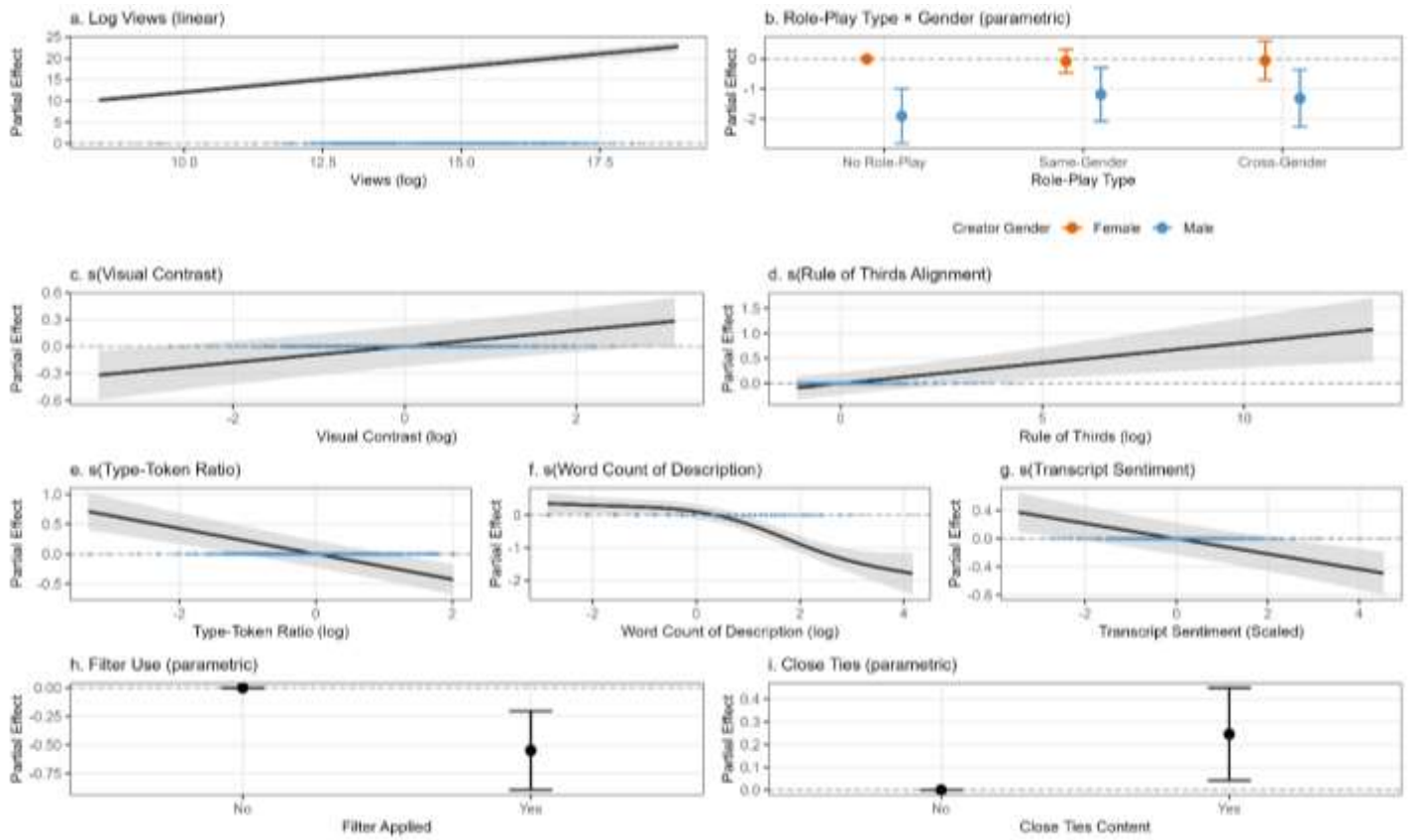
The baseline model (Figure 4; Model 3b, Table 4) demonstrated intermediate explanatory power (adjusted $R^2 = 0.677$, $AIC = 2361.087$). Views showed strong association with shares ($\beta = 1.203$, $SE = 0.034$, $p < .001$). Same-gender role-play significantly predicted elevated shares ($\beta = 0.367$, $SE = 0.135$, $p = .007$). Cross-gender role-

play likewise increased shares ($\beta = 0.410$, $SE = 0.182$, $p = .024$). These corresponded to approximately 44% and 51% more shares respectively. Role-play content travels further through social networks.

Adding role-play type by gender interactions improved fit (adjusted $R^2 = 0.683$, $AIC = 2357.497$). The same-gender role-play by male creator interaction reached significance ($\beta = 0.794$, $SE = 0.271$, $p = .004$). Male creators producing same-gender role-play substantially narrowed the baseline gender gap in shares. The cross-gender interaction was directionally consistent ($\beta = 0.644$, $SE = 0.385$, $p = .095$) but did not reach significance. Several production features shaped sharing behaviour (see Figure 4). Description length produced strong nonlinear effects ($edf = 2.900$, $F = 16.701$, $p < .001$). Greater lexical diversity predicted more shares ($edf = 1.000$, $F = 11.504$, $p < .001$). Positive transcript sentiment also increased sharing ($edf = 1.000$, $F = 5.486$, $p = .019$). Higher visual contrast increased shares ($edf = 1.000$, $F = 4.256$, $p = .039$). Relationship-focused content substantially increased sharing ($\beta = 0.245$, $SE = 0.104$, $p = .018$). Filter use predicted lower shares ($\beta = -0.549$, $SE = 0.177$, $p = .002$). Male creators showed significantly lower baseline sharing ($\beta = -1.910$, $SE = 0.468$, $p < .001$), a pattern absent for likes or comments.

Incorporating role-play-specific production elements (Model 3c, Table 4) degraded fit (adjusted $R^2 = 0.681$, $AIC = 2382.671$). Yet role-play effects persisted (same-gender: $\beta = 0.450$, $SE = 0.152$, $p = .003$; cross-gender: $\beta = 0.441$, $SE = 0.209$, $p = .035$). Relationship-focused content maintained positive association with sharing ($\beta = 0.263$, $SE = 0.105$, $p = .013$). Role-play-specific visual and vocal features offered negligible additional explanatory power.

Figure 26 Significant Smooth and Parametric Terms in Shares Model (3b)



Discussion

This study examined how role-play content differs from non-role-play material in production elements, and how these differences shape audience engagement on TikTok. Three findings stand out. Role-play videos exhibit distinctive multimodal production signatures. Role-play independently predicts higher likes and shares but not comments. Gender-differentiated effects show that male creators benefit most from same-gender role-play in terms of shareability. The sections below interpret these findings against each research question and consider theoretical and practical implications.

5.1 Production Element Differences Between Role-Play and Non-Role-Play Content (RQ1)

Role-play videos systematically differ from non-role-play content across multiple production dimensions, even accounting for creator-level variation. Table 5 summarises these differences ranked by effect magnitude, providing readers with a reference point for evaluating relative importance of each cue. The strongest differentiators are accessories, clothing, close ties content, and physical settings. This pattern aligns with performance theory (Gran, 2025). Role-play necessitates character embodiment, demanding tangible semiotic markers signalling character boundaries to audiences. The production signatures identified here extend earlier qualitative accounts of role-play conventions (Xing, 2024) by providing the first systematic quantitative evidence of their scope and magnitude.

Table 14 Summary of Significant Production Element Differences Between Role-Play Types and Non-Role-Play Content

Predictor	Same-Gender	Cross-Gender	Category
Costume and Styling			
Accessories	OR = 90.46***	OR = 91.54***	Very Strong
Clothing	OR = 4.21***	OR = 36.33***	Strong
Physical settings	OR = 3.05*	—	Moderate
Audio Characteristics			
RMS amplitude (↓)	$\beta = -906.69***$	$\beta = -1178.30***$	Very Strong
Zero crossing rate (↓)	$p < .001***$	—	Strong
Mean pitch modulation	$\beta = 0.83*$ (interaction)	—	Moderate
Background music	OR = 11.49***	—	Strong
Narrative Content			
Close ties content	OR = 7.35***	OR = 26.60***	Very Strong
Platform Features			
Green screen	OR = 20.18**	OR = 28.54***	Very Strong

Filters	OR = 22.70***	OR = 8.32***	Strong
Stitch/duet	OR = 12.67*	—	Moderate
Linguistic Features			
Video duration (interaction)	$\beta = 26.02^{***}$ (male)	Marginal	Moderate
Title sentiment (interaction)	$\beta = -0.48^{**}$	—	Moderate
First-person pronouns (interaction)	$\beta = -0.015^{**}$	—	Moderate
			Non-Significant
Hairstyling	ns	ns	—
Makeup	ns	ns	—

Note: OR = odds ratio; β = standardised coefficient; $\downarrow$ = lower values associated with role-play; * = $p < .05$; ** = $p < .01$; *** = $p < .001$; ns = not significant after FDR correction; — = not tested or not significant; Gender interaction effects indicate differential modulation by creator gender in same-gender role-play contexts.

Costume and Accessories as Primary Markers

Creators employ costume changes and distinctive accessories to differentiate characters. Same-gender role-play showed 4.2 times higher likelihood of distinct clothing compared to non-role-play content. Cross-gender role-play showed 36.3 times higher likelihood. Accessories showed similarly strong associations, with odds ratios exceeding 90 for both role-play types. These visual elements function as essential semiotic markers distinguishing role-play from everyday self-presentation.

This reliance on costume over elaborate makeup aligns with accessible production logic documented on short-form platforms (Xing, 2024). Yet it raises tension with wider literature. Berry et al. (2022) demonstrated that convincing character portrayal relies on coordinated multichannel cues, including facial configuration. Here, creators achieve character legibility through visual shorthand rather than embodied facial transformation. Platform constraints such as short durations and vertical framing privilege instantly readable cues over subtle craft. This supports Foster and Baker's (2022) observation that successful cross-gender clips often depend on instantly recognisable clichés rather than nuanced performance.

The democratising implication is significant. Creators need not invest in professional makeup artistry to construct legible characters. Costume, accessories, and setting changes suffice. This contrasts markedly with classical performance traditions such as Cantonese opera (Tse, 2021), where character embodiment involves elaborate costuming, codified gesture, and carefully managed offstage persona. On TikTok, performance conventions are compressed by platform affordances and time constraints.

Audio Characteristics as Strong Cues

Audio characteristics emerged as particularly robust markers of role-play content. Zero-crossing rate, or audio brightness (the proportion of high-frequency components in audio signals), was significantly higher in same-gender role-play videos. This establishes audio brightness as one of the strongest cues distinguishing role-play from baseline content. Role-play videos employ more varied and dynamic audio landscapes than everyday content.

Beyond audio brightness, mean pitch modulation also differentiated role-play performances. Mean pitch varied significantly by creator gender ($\beta = 0.83$, $SE = 0.40$, $p = .037$ for same-gender role-play), likely reflecting gender-specific vocal affordances. Male voices occupy lower frequency ranges, providing greater scope for pitch variation when portraying different-gender characters. Female creators, with naturally higher fundamental frequencies, face different physiological constraints when differentiating characters vocally. This finding is novel: it demonstrates for the first time how same-gender and cross-gender role-play differ systematically in mean pitch use by creator gender.

This finding extends prior research on audio features in short-form video (Yang et al., 2025). Performance format shapes acoustic as well as visual production choices. The stronger gender interaction for zero-crossing rate indicates that male creators in particular modulate audio when employing role-play. These patterns support Alazhare (2025) and Chen (2024), who found gender-differentiated audio production strategies in short-form video. They resonate with Rakugo literature, where vocal production is tightly bound to gender ideology and comic convention (Şahin, 2022). This study translates that observation into quantitative platform context, revealing gender-specific pitch modulation in role-play adoption for the first time.

Linguistic and Narrative Choices in Role-Play

Gender interactions emerged in linguistic choices across role-play contexts. Male creators produced substantially longer same-gender role-play videos than female creators ($\beta = 26.02$, $SE = 7.36$, $p < .001$). Cross-gender role-play showed a similar though weaker pattern ($\beta = 18.43$, $SE = 10.99$, $p = .094$). Title sentiment revealed gender differentiation ($\beta = -0.48$, $SE = 0.17$, $p = .004$), indicating male creators

employed different affective framings. First-person pronoun usage showed significant interactions ($p = .006$), with male creators reducing self-reference more than female creators in role-play content.

These linguistic patterns suggest that gender shapes not only vocal but also textual strategies in role-play adoption. Male creators may employ longer durations and modified affective framing to establish character legitimacy, potentially reflecting audience expectations about gendered performance.

The association between role-play and close ties content is striking. Both same-gender and cross-gender performances were substantially more likely to address relationship themes than non-role-play content. Role-play functions not merely as comedic technique but as a narrative strategy for depicting interpersonal dynamics efficiently. Rather than describing relationships through monologue, creators use character embodiment to show relational patterns directly. This aligns with sketch comedy traditions prioritising showing over telling (Cooper, 2023).

Cross-gender performances were particularly oriented toward close ties scenarios. Same-gender performance was 7.3 times more likely to address relationship themes ($OR = 7.35$, $FDR\ p < .001$). Cross-gender role-play showed even stronger associations ($OR = 26.60$, $FDR\ p < .001$). This pattern extends Xing's (2024) finding on the influence of couple culture in cross-gender TikTok content. The quantitative evidence here demonstrates this is not incidental: relationship-focused narratives and role-play are structurally connected on this platform. Interpersonal conflict and family dynamics provide culturally legible frameworks for gender contrast. Tang et al. (2021) identified similar reliance on domestic and romantic scenarios in Douyin cross-gender performance. The consistency across Chinese and international platforms suggests narrative choices reflect genre conventions rather than country-specific preferences.

Elements Not Associated with Role-Play

Several production elements showed minimal association with role-play type. Background music, filters, and stitch or duet functions appeared infrequently across all content categories. Low adoption rates suggest creators prioritise authenticity and accessibility over technical sophistication. This aligns with research indicating audiences increasingly demand authentic rather than heavily produced content (Barta & Andalibi,

2021). Effective character embodiment relies primarily on costume, performance, and editing rather than elaborate technical effects.

Title word count showed only marginal differences between role-play and non-role-play content. Cross-gender videos featured notably longer titles, though neither effect survived FDR correction. This may reflect the need to establish narrative context more explicitly when gender boundary crossing occurs, helping audiences interpret performance framing correctly.

5.2 Production Elements, Role-Play Features, and Audience Engagement (RQ2)

Findings addressing RQ2 reveal that engagement is multidimensional rather than unitary. Role-play exerts differential effects across likes, comments, and shares. This supports prior arguments that these metrics reflect qualitatively distinct audience behaviours (Habibi & Salim, 2021; Shahbaznezhad et al., 2021). It has important implications for how researchers and practitioners conceptualise engagement on short-form video platforms, particularly in comedic contexts.

5.2.1 Independent Effects of Role-Play on Engagement Metrics (RQ2a)

Likes

Previous research established that humour and character-driven content increase engagement (Chen, 2024). This study extends that finding substantially. Both same-gender and cross-gender role-play predicted significantly higher like counts relative to non-role-play content, representing approximately 21% increase for each role-play type. The critical novelty lies in the independence of this effect. Chen (2024) did not conduct comprehensive controls for multimodal features. The present findings demonstrate that role-play itself, beyond the costume, accessories, and audio characteristics that accompany it, drives audience approval. Character embodiment generates engagement through distinct mechanisms separate from production feature accumulation.

Likes represent low-threshold approval behaviours requiring minimal cognitive investment (Aldlimi et al., 2025). Role-play content generates approval through entertainment value, narrative complexity, and relatable interpersonal dynamics. The role-play effect holds regardless of creator gender, suggesting that

role-play generates likes through universal mechanisms of entertainment and narrative appeal rather than gender-specific strategies.

Shares

Role-play demonstrated even stronger effects on shares than on likes. Both same-gender role-play (44% increase) and cross-gender role-play (51% increase) significantly elevated share counts relative to non-role-play content. Yet the novel contribution here differs from prior work. Shares require greater investment than likes, involving network distribution and implicit endorsement of content to one's social circle (Achen et al., 2024). Users must consider whether content aligns with their social identity and will resonate with their networks.

Critically, role-play's shareability advantage persists even after controlling for the multimodal production elements accompanying it. Even accounting for costume changes, audio brightness, and relationship-focused narratives, role-play itself drives network amplification. This reveals that character embodiment activates distinct sharing motivations beyond those driven by relational narratives or production choices alone. Role-play provides conversation starters and social currency, enabling viewers to tag others who identify with portrayed relational dynamics (Wahid et al., 2023). Close ties content also significantly increased shares. Thus role-play and relationship-focused content function synergistically for shareability, with their combined effects exceeding simple addition.

Comments

Neither same-gender nor cross-gender role-play significantly predicted comment counts when controlling for production elements. Comments require active discursive participation: viewers must formulate text, expose their thoughts publicly, and potentially engage in dialogue (Habibi & Salim, 2021). Role-play comedy generates approval and sharing without stimulating discussion.

Commenting behaviour responds primarily to controversial, ambiguous, or personally relevant content inviting debate or personal disclosure (Tenenboim, 2022). Role-play comedy typically presents resolved narrative arcs and finished comedic moments. The performance is closed. Cross-age portrayals, by contrast, significantly increased comments, precisely because intergenerational dynamics invoke personal family

experiences and invite evaluative responses. This distinction illuminates a boundary: content that entertains differs fundamentally from content that provokes. Role-play operates primarily in the entertainment domain, whilst cross-generational scenarios activate evaluative engagement.

5.2.2 Cross-Gender Role-Play and Gender-Differentiated Engagement Effects (RQ2b)

Gender Performance and Engagement

Several findings stand in productive tension with the reviewed literature. Wu et al. (2025) found that gender-incongruent content generated higher engagement for chief executive officers on microblogs, attributing this to novelty and expectation violation. The present study found weaker evidence for this mechanism in TikTok comedic content. Cross-gender role-play showed only trends toward higher shares and failed to reach significance for other metrics.

One explanation is normalisation. TikTok's prevalence of cross-gender performance may have reduced the novelty driving expectation-violation effects. Cross-gender role-play is now an established genre on TikTok, appearing in approximately 12.5% of sampled videos. Where Wu et al. examined a professional context where gender incongruence remains relatively rare, TikTok audiences have developed genre expectations that blunt the novelty response. The mechanism may not transfer across contexts. This interpretation aligns with Xing's (2024) observation that cross-gender practices on TikTok partially deconstruct gender binaries without overturning entrenched norms. The platform rewards highly legible gender performances, and audiences appear to have developed genre expectations blunting novelty effects. Algorithmic bias documented by Fosch-Villaronga et al. (2021) and Gerrard and Thornham (2020) may compound this effect, constraining non-conforming performances independently of intrinsic appeal.

The near-absence of prosthetic makeup aligns with and extends Tang et al.'s (2021) critical discourse findings. Tang et al. found that Douyin cross-gender performances rely on recurrent gestural and narrative tropes reinforcing gender stereotypes. This study finds parallel reliance on accessible visual shorthand over subtle embodied transformation. Both suggest that cross-gender performance on short-form platforms tends toward convention rather than subversion. Whether this reflects platform affordances, audience expectations, or algorithmic incentives remains an open question.

Gender Interactions for Shares

The interaction model for shares improved fit and was selected based on AIC comparison. Male creators experienced substantial sharing gains from same-gender role-play ($\beta = 0.794$, $SE = 0.271$, $p = .004$). Female creators showed minimal role-play effects on sharing. Yet male creators showed significantly lower baseline sharing across all content types.

Character embodiment may constitute gender-atypical behaviour for male creators. Emotional expressiveness and character performance traditionally align with feminine performance norms. This boundary crossing enhances novelty and appeal, particularly for male creators whose typical content relies on observational or physical comedy. Critically, this effect persists even after controlling for the multimodal production elements accompanying role-play. The robust sharing gain for male creators reflects the performance mode itself, not differences in costume use, audio modulation, or narrative content. This finding resonates with Wu et al.'s (2025) expectation-violation mechanism, applied here at the level of performance mode rather than content topic. It aligns with Wan and Lu's (2024) finding that creators crossing normative gender boundaries foster distinctive audience engagement.

Cross-Gender Role-Play Effects

The interaction model for shares improved fit and was selected based on AIC comparison. Male creators experienced substantial sharing gains from same-gender role-play ($\beta = 0.794$, $SE = 0.271$, $p = .004$). Female creators showed minimal role-play effects on sharing. Yet male creators showed significantly lower baseline sharing across all content types.

Character embodiment may constitute gender-atypical behaviour for male creators. Emotional expressiveness and character performance traditionally align with feminine performance norms. This boundary crossing enhances novelty and appeal, particularly for male creators whose typical content relies on observational or physical comedy. Critically, this effect persists even after controlling for the multimodal production elements accompanying role-play. The robust sharing gain for male creators reflects the performance mode itself, not differences in costume use, audio modulation, or narrative content. This finding resonates with Wu et al.'s (2025) expectation-violation mechanism, applied here at the level of

performance mode rather than content topic. It aligns with Wan and Lu's (2024) finding that creators crossing normative gender boundaries foster distinctive audience engagement.

Production Strategy Differences by Creator Gender

Male creators employing same-gender role-play produced longer videos, modulated pitch differently, and incorporated background music more frequently than female creators. These patterns suggest gender-specific performance strategies shaped by different audience expectations and physiological affordances. Male creators may require longer durations to establish character boundaries convincingly, whilst female creators achieve comparable effects more efficiently. The stronger pitch modulation effects for male creators align with vocal affordance arguments in Section 2.3 and with the multimodal coordination findings of Berry et al. (2022). Together, these patterns suggest that gender-specific production strategies emerge from interaction between physiological affordances, platform conventions, and audience expectations.

5.3 Theoretical and Practical Implications

Theoretical Implications

The findings advance performance theory in four respects. First, they demonstrate that Goffman's (1959) dramaturgical framework, whilst valuable for conceptualising online self-presentation, requires extension to accommodate explicit character performance. Role-play represents deliberate theatrical embodiment distinct from everyday persona construction (Gran, 2025). The production signatures distinguishing role-play signal this theatrical framing, enabling audiences to interpret content as entertainment rather than authentic self-disclosure.

Second, results support multimodal communication theories emphasising that engagement emerges from coordinated visual, auditory, and linguistic features. Critically, these effects differ systematically by creator gender. The findings reveal complex, nonlinear relationships. Description length showed diminishing returns at higher word counts for likes, comments, and shares. Affective valence produced an inverted-U relationship with comments, indicating that moderately valenced content generated more discussion than either neutral or extreme material. These nonlinear patterns, alongside role-play type by gender interactions, underscore the contingent nature of engagement determinants. Simple additive models cannot capture these

dynamics. Generalised additive mixed models prove necessary for detecting these interaction patterns and nonlinearities. The strong audio effects, combined with costume and accessory markers, demonstrate that role-play depends on synchronised multimodal cues to signal character boundaries.

Third, the differential effects across engagement types advance understanding of engagement as a multidimensional construct. Although likes, shares, and comments are highly intercorrelated, they involve different degrees of active participation and social identity considerations. Researchers must specify which engagement behaviours they examine and avoid conflating metrics reflecting fundamentally different processes.

Fourth, results illuminate how platform affordances shape creative labour. TikTok's algorithmic prioritisation of engagement metrics creates incentives for role-play adoption even amongst creators with modest budgets. Creative skill matters more than production infrastructure. Yet this simultaneously intensifies creative labour demands, requiring continuous performance innovation to maintain algorithmic visibility.

Practical Implications

Creators seeking to enhance engagement should consider role-play strategies, particularly when addressing relational themes. Multi-character performances depicting family dynamics, friendships, or romantic scenarios demonstrate strong engagement potential. Effective character differentiation can be achieved through costume changes, accessories, and vocal modulation. Elaborate production infrastructure is not required.

Male creators may particularly benefit from role-play adoption given the observed interaction effects. Character embodiment appears to help male creators overcome lower baseline sharing through performance strategies enhancing novelty and audience connection. Same-gender role-play benefits male creators more than female creators, with direct implications for content strategy.

Platform designers should recognise that engagement optimisation requires attending to differential effects across engagement types. Algorithmic systems weighting likes, comments, and shares equally may

misallocate visibility to content excelling on single dimensions. Metric-specific weighting reflecting different user investment levels could improve both content quality and user satisfaction.

Communication practitioners developing branded content or influencer partnerships should recognise that role-play formats enhance shareability more than commenting behaviour. Campaigns prioritising word-of-mouth amplification may benefit from role-play approaches featuring relatable interpersonal scenarios.

Campaigns seeking discursive engagement require alternative strategies such as open-ended questions or calls for personal anecdotes.

5.4 Strengths, Limitations, and Future Directions

Methodological Strengths

This study offers several methodological strengths. The comprehensive multimodal feature set spans visual, acoustic, and linguistic dimensions, enabling holistic characterisation of role-play content compared to studies examining isolated cues. Treating engagement as a multidimensional construct with separate models for likes, comments, and shares reveals patterns that unidimensional approaches would obscure. Generalised additive mixed models accommodate both nonlinear relationships and hierarchical data structure, improving on fixed-effects and single-metric approaches dominating prior work.

Coding of binary production features involved trained researchers calibrated against platform affordances, ensuring accuracy for TikTok-specific features. Inter-coder reliability checks confirmed consistency across coders, validating manual annotation accuracy. This approach balances rigour with the practical limitations of automated detection for nuanced production choices.

Limitations and Future Directions

The sample comprised high-profile creators active during the COVID-19 pandemic. The pandemic period was characterised by elevated platform use and heightened appetite for escapist content. It remains to be seen whether role-play's advantages persist post-pandemic. Likewise, it remains unclear whether findings apply to mid-tier or micro-creators, given different follower dynamics and algorithmic treatment. The focus on comedic content further limits generalisability. Role-play may function differently in informational, educational, or advocacy genres serving distinct communicative purposes.

The cross-gender video subsample was relatively small ($n = 60$). Effects for cross-gender role-play showed clear directional patterns and statistical significance for overall role-play status, yet a larger sample would strengthen conclusions about cross-gender-specific effects. Replication with a larger and more balanced sample would provide additional confidence in cross-gender engagement patterns.

The role of artificial intelligence in short-form video production warrants particular attention (Pellas, 2023). AI-powered editing tools, voice modulation software, and generative image technologies are increasingly accessible to creators (Banh & Strobel, 2023). These tools could substantially alter the production signatures currently distinguishing role-play content. AI-generated voices and automated scene switching may lower production barriers further (Anantrasirichai et al., 2025). Yet they introduce performance dimensions that existing coding schemes do not capture (Wen et al., 2020; Xu et al., 2025). Whether AI-assisted production amplifies or disrupts multimodal coherence remains unclear. Future research should examine how recommendation systems interact with AI-assisted role-play content and whether engagement mechanisms documented here remain stable as production practices evolve.

Conclusion

This study examined which production elements distinguish role-play from non-role-play comedic content on TikTok. It investigated how role-play and those elements predict audience engagement. Three findings are central. First, role-play videos exhibit distinctive multimodal production signatures. These include costume, accessories, audio brightness and pitch modulation with gender-differentiated effects. They also include relationship-focused narratives and linguistic features such as duration, title sentiment, and first-person pronoun usage. Second, role-play independently predicts higher likes and shares but not comments. This reveals engagement as multidimensional rather than unitary. Third, gender-differentiated effects show that same-gender role-play benefits male creators disproportionately in shareability. This pattern aligns with expectation-violation accounts of gender-atypical performance (Wu et al., 2025).

The study makes four contributions. First, it provides the first systematic quantitative characterisation of role-play production signatures. It demonstrates that costume, accessories, audio modulation, and linguistic modulation function as functional markers of character embodiment. These are distinct from mere aesthetic

choices (Gran, 2025; Xing, 2024). Mean pitch variation shows particularly important gender-differentiated effects reflecting physiological affordances. Second, it establishes that role-play exerts independent engagement effects. These persist even after controlling comprehensively for production elements and creator-level variation. This independent effect extends prior work (Chen, 2024) by isolating role-play's contribution beyond correlated multimodal features.

It advances understanding of engagement as multidimensional. Likes, shares, and comments respond to qualitatively distinct audience motivations (Habibi & Salim, 2021; Shahbaznezhad et al., 2021). Description length showed diminishing returns at higher word counts. Transcript sentiment produced an inverted-U relationship with comments. This demonstrates the necessity of nonlinear analytical approaches. Generalised additive mixed models accommodate nonlinear relationships and hierarchical data structure. They are necessary for detecting interaction patterns that fixed-effects approaches would obscure.

These findings carry practical weight. Creators can achieve effective character differentiation through costume, accessories, and vocal modulation. Elaborate production infrastructure is not required. Male creators may particularly benefit from role-play adoption. Platform designers should recognise that systems weighting all engagement types equally may misallocate visibility. The platform environment rewards legible gender performances over experimental ones (Foster & Baker, 2022; Xing, 2024). This has implications for content diversity amplified by recommendation systems.

As short-form platforms evolve, understanding how performance characteristics interact with affordances becomes essential. This applies for creators, platforms, and audiences navigating algorithmically mediated environments.

References (Chapter 3):

- Achen, R. M., Stadler-Blank, A., & Sailors, J. J. (2024). I “like” it: The effects of social media platform and message on consumer engagement actions. *International Journal of Sport Communication*, 17(1), 5–16. <https://doi.org/10.1123/ijsc.2023-0125>
- Ahlse, J., Nilsson, F., & Sandström, N. (2020). *It's time to TikTok: Exploring Generation Z's motivations to participate in #Challenges*. <https://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-48708>
- Åkestam, N., Rosengren, S., Dahlén, M., Liljedal, K. T., & Berg, H. (2021). Gender stereotypes in advertising have negative cross-gender effects. *European Journal of Marketing*, 55(13), 63–93. <https://doi.org/10.1108/EJM-02-2019-0125>
- Aldlimi, T., Priporas, C., & Chang, S. (2025). Engagement in influencer marketing: A systematic review of key drivers, behaviors, and future research directions. *Journal of Consumer Behaviour*, 24(3), 1541–1566. <https://doi.org/10.1002/cb.2485>
- Anantrasirichai, N., Zhang, F., & Bull, D. (2025). *Advances in artificial intelligence: A review for the creative industries*. arXiv. <https://doi.org/10.48550/ARXIV.2501.02725>
- Ao, L., Bansal, R., Pruthi, N., & Khaskheli, M. B. (2023). Impact of social media influencers on customer engagement and purchase intention: A meta-analysis. *Sustainability*, 15(3), 2744. <https://doi.org/10.3390/su15032744>
- Ashraf, A., Hameed, I., & Saeed, S. A. (2023). How do social media influencers inspire consumers' purchase decisions? The mediating role of parasocial relationships. *International Journal of Consumer Studies*, 47(4), 1416–1433. <https://doi.org/10.1111/ijcs.12917>
- Ballester, E., Ruiz, C., Rubio, N., & Veloutsou, C. (2025). We match! Building online brand engagement behaviours through emotional and rational processes. *Journal of Retailing and Consumer Services*, 82, 104146. <https://doi.org/10.1016/j.jretconser.2024.104146>
- Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, 33(1), 63. <https://doi.org/10.1007/s12525-023-00680-1>

- Barrie, C., Atalay, A., & ElKattan, A. (2025). *Information diets are more diverse in attention than in engagement*. https://doi.org/10.31219/osf.io/kvbzd_v3
- Barta, K., & Andalibi, N. (2021). Constructing authenticity on tiktok: Social norms and social support on the ‘fun’ platform. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–29. <https://doi.org/10.1145/3479574>
- Barthelme, S. (2015). *Imager: Image processing library based on ‘cimg’* [Data set]. <https://doi.org/10.32614/CRAN.package.imager>
- Ben-Iheanacho, E. O. (2024). Gender swap, cross-dressing, and androgyny in Oyin Olugbile’s Sanya. *Journal of Languages, Linguistics and Literary Studies*, 4(3), 131–138. <https://doi.org/10.57040/3dqkka72>
- Berry, M., Lewin, S., & Brown, S. (2022). Correlated expression of the body, face, and voice during character portrayal in actors. *Scientific Reports*, 12(1), 8253. <https://doi.org/10.1038/s41598-022-12184-7>
- Bhandari, A., & Bimo, S. (2022). Why’s everyone on tiktok now? The algorithmized self and the future of self-making on social media. *Social Media + Society*, 8(1), 20563051221086241. <https://doi.org/10.1177/20563051221086241>
- Bilozerskyi, V., Dergachov, K., Krasnov, L., Zymovin, A., & Popov, A. (2023). New method for video stream brightness stabilization: Algorithms and performance evaluation. *Radioelectronic and Computer Systems*, (3), 125–135. <https://doi.org/10.32620/reks.2023.3.10>
- Boeker, M., & Urman, A. (2022). An empirical investigation of personalization factors on TikTok. *Proceedings of the ACM Web Conference 2022*, 2298–2309. <https://doi.org/10.1145/3485447.3512102>
- Bouchaud, P. (2024). Algorithmic amplification of politics and engagement maximization on social media. In H. Cherifi, L. M. Rocha, C. Cherifi, & M. Donduran (Eds), *Complex Networks & Their Applications XII* (Vol. 1144, pp. 131–142). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-53503-1_11

- Brush, H. E. (2024). *Cosplay and the Impact on Body Image* [Doctoral Dissertation, University of St. Thomas]. <https://researchonline.stthomas.edu/esploro/outputs/doctoral/Cosplay-and-the-Impact-on-Body/991015207701103691#file-0>
- Carter, A., & Davies, K. (2025). Laughing through Brexit: Family humour practices, political troubles and everyday life. *Current Sociology*, 73(7), 1111–1128. <https://doi.org/10.1177/00113921241275684>
- Cervi, L., & Divon, T. (2023). Playful Activism: Memetic Performances of Palestinian Resistance in TikTok #Challenges. *Social Media + Society*, 9(1), 205630512311576. <https://doi.org/10.1177/20563051231157607>
- Chan, T. H., Hung, K., & Tse, D. K. (2023). Comparing e-commerce micro- and macroinfluencers in tiktok videos: Effects of strategies on audience likes, audience shares, and brand sales. *Journal of Interactive Advertising*, 23(4), 307–322. <https://doi.org/10.1080/15252019.2023.2273253>
- Charlesworth, T. E. S., Yang, V., Mann, T. C., Kurdi, B., & Banaji, M. R. (2021). Gender stereotypes in natural language: Word embeddings show robust consistency across child and adult language corpora of more than 65 million words. *Psychological Science*, 32(2), 218–240. <https://doi.org/10.1177/0956797620963619>
- Chen, J. (2024). Application of audiovisual language on the TikTok platform and analysis of its communication effect. *Advances in Social Behavior Research*, 11, 33–40. <https://doi.org/10.54254/2753-7102/11/2024104>
- Chen, Q., Min, C., Zhang, W., Ma, X., & Evans, R. (2021). Factors driving citizen engagement with government tiktok accounts during the covid-19 pandemic: Model development and analysis. *Journal of Medical Internet Research*, 23(2), e21463. <https://doi.org/10.2196/21463>
- Cheng, Z., & Li, Y. (2024). Like, comment, and share on tiktok: Exploring the effect of sentiment and second-person view on the user engagement with tiktok news videos. *Social Science Computer Review*, 42(1), 201–223. <https://doi.org/10.1177/08944393231178603>

- Chung, J. (Jae), Ding, Y., & Kalra, A. (2023). I really know you: How influencers can increase audience engagement by referencing their close social ties. *Journal of Consumer Research*, 50(4), 683–703. <https://doi.org/10.1093/jcr/ucad019>
- Cook, C., Callahan, J. L., Pollet, T. V., & Elliott, C. (2024). Gender(Ed) performances: Women's impression management in stand-up comedy. *Human Relations*, 77(4), 533–559. <https://doi.org/10.1177/00187267221137996>
- Cooper, E. (2023). TikTok and the Romcom: How User-Generated Content Allows Audiences to Reclaim Traditional Film Tropes. *FMST 001 Critical Approaches to Media*, 1(2), 34–42. <https://www.swarthmore.edu/sites/default/files/assets/documents/film-media-studies/FMST%20001%20Journal%20%232%20Social%20Media%207.pdf#page=36>
- Crigler, R. K. (2024). Afropolitan influence: Gender, comedy, and social media in global africa. *Social Media + Society*, 10(4), 20563051241308330. <https://doi.org/10.1177/20563051241308330>
- Dong, X., Liu, H., Xi, N., Liao, J., & Yang, Z. (2024). Short video marketing: What, when and how short-branded videos facilitate consumer engagement. *Internet Research*, 34(3), 1104–1128. <https://doi.org/10.1108/INTR-02-2022-0121>
- Duffy, B. E., Pinch, A., Sannon, S., & Sawey, M. (2021). The nested precarities of creative labor on social media. *Social Media + Society*, 7(2), 20563051211021368. <https://doi.org/10.1177/20563051211021368>
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. <https://doi.org/10.1037/0033-295X.109.3.573>
- Er, M. B. (2020). A novel approach for classification of speech emotions based on deep and acoustic features. *IEEE Access*, 8, 221640–221653. <https://doi.org/10.1109/ACCESS.2020.3043201>
- Firth, E., & Marinelli, A. (2025). Datacasting: Tiktok's algorithmic flow as televisual experience. *Media and Communication*, 13, 9392. <https://doi.org/10.17645/mac.9392>

- Fosch-Villaronga, E., Poulsen, A., Søraa, R. A., & Custers, B. H. M. (2021). A little bird told me your gender: Gender inferences in social media. *Information Processing & Management*, 58(3), 102541. <https://doi.org/10.1016/j.ipm.2021.102541>
- Gerrard, Y., & Thornham, H. (2020). Content moderation: Social media's sexist assemblages. *New Media & Society*, 22(7), 1266–1286. <https://doi.org/10.1177/1461444820912540>
- Gerwin, J. (2024). The role of dialect in comedy performances: Focus on enregisterment and humor. *English Language and Linguistics*, 28(4), 751–773. <https://doi.org/10.1017/S1360674324000492>
- Gran, A.-B. (2025). Performing not-not-me in some: A new theatrical typology of self-presentation online. *Social Media + Society*, 11(1), 20563051251315256. <https://doi.org/10.1177/20563051251315256>
- Guinaudeau, B., Munger, K., & Votta, F. (2022). Fifteen Seconds of Fame: TikTok and the Supply Side of Social Video. *Computational Communication Research*, 4(2), 463–485. <https://doi.org/10.5117/CCR2022.2.004.GUIN>
- Habibi, S. A., & Salim, L. (2021). Static vs. dynamic methods of delivery for science communication: A critical analysis of user engagement with science on social media. *PLOS ONE*, 16(3), e0248507. <https://doi.org/10.1371/journal.pone.0248507>
- Han, Y., & Ho, X. (2024). “More than just playing a character”: Gender exploration and expression in avatar customization. *Games and Culture*, 15554120241307982. <https://doi.org/10.1177/15554120241307982>
- Hautea, S., Parks, P., Takahashi, B., & Zeng, J. (2021). Showing they care (Or Don't): Affective publics and ambivalent climate activism on tiktok. *Social Media + Society*, 7(2), 20563051211012344. <https://doi.org/10.1177/20563051211012344>
- Herrick, S. S. C., Hallward, L., & Duncan, L. R. (2021). “This is just how I cope”: An inductive thematic analysis of eating disorder recovery content created and shared on TikTok using # EDrecovery. *International Journal of Eating Disorders*, 54(4), 516–526. <https://doi.org/10.1002/eat.23463>

- Hove, M. J., Vuust, P., & Stupacher, J. (2019). Increased levels of bass in popular music recordings 1955–2016 and their relation to loudness. *The Journal of the Acoustical Society of America*, 145(4), 2247–2253. <https://doi.org/10.1121/1.5097587>
- Hudders, L., & De Jans, S. (2022). Gender effects in influencer marketing: An experimental study on the efficacy of endorsements by same- vs. other-gender social media influencers on Instagram. *International Journal of Advertising*, 41(1), 128–149. <https://doi.org/10.1080/02650487.2021.1997455>
- Hussain, K., Khan, M. L., & Malik, A. (2024). Exploring audience engagement with ChatGPT-related content on YouTube: Implications for content creators and AI tool developers. *Digital Business*, 4(1), 100071. <https://doi.org/10.1016/j.digbus.2023.100071>
- Jaakonmäki, R., Müller, O., & Vom Brocke, J. (2017). *The impact of content, context, and creator on user engagement in social media marketing*. Hawaii International Conference on System Sciences. <https://doi.org/10.24251/HICSS.2017.136>
- Jin, X., Li, X., Lou, H., Fan, C., Deng, Q., Xiao, C., Cui, S., & Singh, A. K. (2022). Aesthetic attribute assessment of images numerically on mixed multi-attribute datasets. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 18(3s), 1–16. <https://doi.org/10.1145/3547144>
- Jockers, M. (2015). *Syuzhet: Extracts sentiment and sentiment-derived plot arcs from text* [Data set]. <https://doi.org/10.32614/CRAN.package.syuzhet>
- Kaye, D. B. V., Chen, X., & Zeng, J. (2021). The co-evolution of two Chinese mobile short video apps: Parallel platformization of Douyin and TikTok. *Mobile Media & Communication*, 9(2), 229–253. <https://doi.org/10.1177/2050157920952120>
- Klug, D. (2020). 'It took me almost 30 minutes to practice this'. *Performance and Production Practices in Dance Challenge Videos on TikTok*. <https://doi.org/10.48550/ARXIV.2008.13040>

- Klug, D., Evans, M., & Kaufman, G. (2023). How TikTok served as a platform for young people to share and cope with lived COVID-19 experiences. *MedieKultur: Journal of Media and Communication Research*, 38(73), 152–170. <https://doi.org/10.7146/mk.v38i73.128463>
- Kostyk, A., & Huhmann, B. A. (2021). Perfect social media image posts: Symmetry and contrast influence consumer response. *European Journal of Marketing*, 55(6), 1747–1779. <https://doi.org/10.1108/EJM-09-2018-0629>
- Kuo, H.-Y., & Chen, S.-Y. (2025). Predicting user engagement in health misinformation correction on social media platforms in taiwan: Content analysis and text mining study. *Journal of Medical Internet Research*, 27, e65631. <https://doi.org/10.2196/65631>
- Leech, G. N., Rayson, P., & Wilson, A. (2014). *Word frequencies in written and spoken english: Based on the british national corpus*. Taylor and Francis. https://ucrel.lancs.ac.uk/bncfreq/lists/2_1_spokenvwritten_alpha.txt
- Li, D., & Guo, J. (2025). Online feminist stand-up comedy: An emerging affective-discursive intervention on gender in China. *Continuum*, 39(1), 34–46. <https://doi.org/10.1080/10304312.2024.2354247>
- Li, Y., Guan, M., Hammond, P., & Berrey, L. E. (2021). Communicating COVID-19 information on TikTok: A content analysis of TikTok videos from official accounts featured in the COVID-19 information hub. *Health Education Research*, 36(3), 261–271. <https://doi.org/10.1093/her/cyab010>
- Ligges, U., Krey, S., Mersmann, O., Schnackenberg, S., Guénard, G., & Ellis, D. P. W. (2004). *Tuner: Analysis of music and speech* [Data set]. <https://doi.org/10.32614/CRAN.package.tuneR>
- Lin, J., Li, W., Zhu, L., Li, N., & Chang, S. (2025). The most popular videos promoting breast enhancement products on tiktok: Cross-sectional content and user engagement analysis. *Journal of Medical Internet Research*, 27, e73336–e73336. <https://doi.org/10.2196/73336>
- Liu, C., Jiang, M., & Muhammad, Z. A. (2024). The impact of TikTok short video factors on tourists' behavioral intention among Generation Z and Millennials: The role of flow experience. *PLOS ONE*, 19(12), e0315140. <https://doi.org/10.1371/journal.pone.0315140>

- Liu, M., Wang, G., & Zhao, R. (2025). “Double Bind” with a twist: A corpus-assisted discourse study of gender performances of male and female entrepreneurs on Twitter (Now x). *PLOS One*, 20(8), e0331400. <https://doi.org/10.1371/journal.pone.0331400>
- Lockyer, S., & De Benedictis, S. (2023). Performing pregnancy: Comic content, critique and ambivalence in pregnant stand-up comedy. *International Journal of Cultural Studies*, 26(3), 343–361. <https://doi.org/10.1177/13678779231166444>
- Logi, L. (2025). War(N/m)ing the room: The role of vocal sound semiotics in nuancing comedian-audience relations. *Journal of World Languages*. <https://doi.org/10.1515/jwl-2025-0038>
- Lu, Y., & (Cindy) Shen, C. (2023). Unpacking multimodal fact-checking: Features and engagement of fact-checking videos on Chinese TikTok (Douyin). *Social Media + Society*, 9(1), 20563051221150406. <https://doi.org/10.1177/20563051221150406>
- Lüdecke, D., Makowski, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., Wiernik, B. M., & Thériault, R. (2019). *Performance: Assessment of regression models performance* [Data set]. <https://doi.org/10.32614/CRAN.package.performance>
- Martínez-López, F. J., Anaya-Sánchez, R., Esteban-Millat, I., Torrez-Meruvia, H., D’Alessandro, S., & Miles, M. (2020). Influencer marketing: Brand control, commercial orientation and post credibility. *Journal of Marketing Management*, 36(17–18), 1805–1831. <https://doi.org/10.1080/0267257X.2020.1806906>
- Mfondoum, A. H. N., Nguet, P. W., Seuui, D. T., Mfondoum, J. V. M., Ngenyam, H. B., Diba, I., Tchindjang, M., Djiangoue, B., Mihi, A., Hakdaoui, S., Batcha, R., Tchatchouang, F. C. L., Petcheu, I. C. N., & Beni, L. M. (2023). Stepwise integration of analytical hierarchy process with machine learning algorithms for landslide, gully erosion and flash flood susceptibility mapping over the North-Moungo perimeter, Cameroon. *Geoenvironmental Disasters*, 10(1), 22. <https://doi.org/10.1186/s40677-023-00254-5>

- Morgenroth, T., & Ryan, M. K. (2021). The effects of gender trouble: An integrative theoretical framework of the perpetuation and disruption of the gender/sex binary. *Perspectives on Psychological Science*, 16(6), 1113–1142. <https://doi.org/10.1177/1745691620902442>
- Morgenroth, T., & Ryan, M. K. (2021). The effects of gender trouble: An integrative theoretical framework of the perpetuation and disruption of the gender/sex binary. *Perspectives on Psychological Science*, 16(6), 1113–1142. <https://doi.org/10.1177/1745691620902442>
- Mullen, L. (2016). *Tokenizers: Fast, consistent tokenization of natural language text* [Data set]. <https://doi.org/10.32614/CRAN.package.tokenizers>
- Munaro, A. C., Hübner Barcelos, R., Francisco Maffezzoli, E. C., Santos Rodrigues, J. P., & Cabrera Paraiso, E. (2021). To engage or not engage? The features of video content on YouTube affecting digital consumer engagement. *Journal of Consumer Behaviour*, 20(5), 1336–1352. <https://doi.org/10.1002/cb.1939>
- Ngilangil, K. M. (2022). Tiktok on SNSU students: Engagement and influence. *International Journal of English Literature and Social Sciences*, 7(4), 150–155. <https://doi.org/10.22161/ijels.74.22>
- O'Rourke, C. (2020). Exploiting ambiguity: *Murder!* and the meanings of cross-dressing in interwar british cinema. *Journal of British Cinema and Television*, 17(3), 289–312. <https://doi.org/10.3366/jbctv.2020.0530>
- Ooms, J. (2016). *Magick: Advanced graphics and image-processing in r* [Data set]. <https://doi.org/10.32614/CRAN.package.magick>
- Osisanwo, A., Atoloye, L., & Akintaro, S. (2024). Pragmatic acts of humour in family discourse in selected Maryam Apaokagi's comedy skits. *The European Journal of Humour Research*, 12(2), 53–73. <https://doi.org/10.7592/EJHR.2024.12.2.843>
- Pan, M., Blut, M., Ghiassaleh, A., & Lee, Z. W. Y. (2025). Influencer marketing effectiveness: A meta-analytic review. *Journal of the Academy of Marketing Science*, 53(1), 52–78. <https://doi.org/10.1007/s11747-024-01052-7>

- Parvin, E. (2024). *Analysis and critique of character creators in role-playing video games towards a more inclusive gaming environment* [MS, Illinois State University].
<https://doi.org/10.30707/ETD2024.20240618063950371920.999940>
- Pellas, N. (2023). The influence of sociodemographic factors on students' attitudes toward AI-generated video content creation. *Smart Learning Environments*, 10(1), 57. <https://doi.org/10.1186/s40561-023-00276-4>
- Perugia, G., & Lisy, D. (2023). Robot's gendering trouble: A scoping review of gendering humanoid robots and its effects on hri. *International Journal of Social Robotics*, 15(11), 1725–1753.
<https://doi.org/10.1007/s12369-023-01061-6>
- Pfeuffer, A., & Huh, J. (2021). Effects of different sponsorship disclosure message types on consumers' trust and attitudes. *International Journal of Advertising*, 40(1), 49–80.
<https://doi.org/10.1080/02650487.2020.1807229>
- Pradhan, B., Kishore, K., & Gokhale, N. (2023). Social media influencers and consumer engagement: A review and future research agenda. *International Journal of Consumer Studies*, 47(6), 2106–2130.
<https://doi.org/10.1111/ijcs.12901>
- Puterbaugh, B., & Medford, K. (2025). The comedic performance of Mock Haitian in Dominican media: A raciosemiotic approach to enregisterment. *Journal of Linguistic Anthropology*, 35(3), e70027.
<https://doi.org/10.1111/jola.70027>
- Qu, T. (2020, April 30). *TikTok and China version Douyin surpass 2 billion download milestone*. South China Morning Post. <https://www.scmp.com/tech/apps-social/article/3082285/tiktok-and-china-version-douyin-surpass-2-billion-download>
- Rejer, I., Jankowski, J., Dreger, J., & Lorenz, K. (2024). Viewer engagement in response to mixed and uniform emotional content in marketing videos—An electroencephalographic study. *Sensors*, 24(2), 517. <https://doi.org/10.3390/s24020517>
- Rinker, T. (2017). *Textstem: Tools for stemming and lemmatizing text* [Data set].
<https://doi.org/10.32614/CRAN.package.textstem>

- Şahin, E. (2022). His master's voice, her jokes: Voice and gender politics in the performance of rakugo. *Journal of Linguistic Anthropology*, 32(3), 476–495. <https://doi.org/10.1111/jola.12351>
- Schellewald, A. (2022). Theorizing “stories about algorithms” as a mechanism in the formation and maintenance of algorithmic imaginaries. *Social Media + Society*, 8(1), 20563051221077025. <https://doi.org/10.1177/20563051221077025>
- Schwenzow, J., Hartmann, J., Schikowsky, A., & Heitmann, M. (2021). Understanding videos at scale: How to extract insights for business research. *Journal of Business Research*, 123, 367–379. <https://doi.org/10.1016/j.jbusres.2020.09.059>
- Shahbaznezhad, H., Dolan, R., & Rashidirad, M. (2021). The role of social media content format and platform in users' engagement behavior. *Journal of Interactive Marketing*, 53(1), 47–65. <https://doi.org/10.1016/j.intmar.2020.05.001>
- Simpson, E., Hamann, A., & Semaan, B. (2022). How to tame ‘your’ algorithm: Lgbtq+ users’ domestication of tiktok. *Proceedings of the ACM on Human-Computer Interaction*, 6(GROUP), 1–27. <https://doi.org/10.1145/3492841>
- Skorodihina, A. (2024). *Gayshin Impact: The LGBTQ+ Gaming Fandom and Queering Practices on Social Media* [BATHesis, Tilburg University]. <http://arno.uvt.nl/show.cgi?fid=178034>
- Smith, J. K., & Mendelson, E. A. (2025). Parasocial parenting, adoption, and monetization of the “internet parent” with the Griswolds on TikTok. *International Journal of Cultural Studies*, 28(5), 907–926. <https://doi.org/10.1177/13678779241288619>
- Su, C., Zhou, H., Gong, L., Teng, B., Geng, F., & Hu, Y. (2021). Viewing personalized video clips recommended by TikTok activates default mode network and ventral tegmental area. *NeuroImage*, 237, 118136. <https://doi.org/10.1016/j.neuroimage.2021.118136>
- Sueur, J., Aubin, T., & Simonis, C. (2006). *Seewave: Sound analysis and synthesis* [Data set]. <https://doi.org/10.32614/CRAN.package.seewave>

- Suhertina, S., Hasbullah, H., Ahmad, R. R. M. R., & Hasgimianti, H. (2025). Exploring muslim youth perceptions of gender imitation and transwoman representation in digital media. *Indonesian Journal of Counseling and Development*, 7(2), 310–329. <https://doi.org/10.32939/ijcd.v7i2.6200>
- Tang, X., Zou, W., Hu, Z., & Tang, L. (2021). The recreation of gender stereotypes in male cross-dressing performances on Douyin. *Journal of Broadcasting & Electronic Media*, 65(5), 660–678. <https://doi.org/10.1080/08838151.2021.1955888>
- Tenenboim, O. (2022). Comments, shares, or likes: What makes news posts engaging in different ways. *Social Media + Society*, 8(4), 20563051221130282. <https://doi.org/10.1177/20563051221130282>
- Thelwall, M., & Foster, D. (2021). Male or female gender-polarized YouTube videos are less viewed. *Journal of the Association for Information Science and Technology*, 72(12), 1545–1557. <https://doi.org/10.1002/asi.24529>
- Trunfio, M., & Rossi, S. (2021). Conceptualising and measuring social media engagement: A systematic literature review. *Italian Journal of Marketing*, 2021(3), 267–292. <https://doi.org/10.1007/s43039-021-00035-8>
- Tse, P. (2021). Performing androgyny: Cross-dressing actresses, fandom, and queer sensibility in Hong Kong Cantonese opera. *Inter-Asia Cultural Studies*, 22(2), 139–157. <https://doi.org/10.1080/14649373.2021.1929403>
- Venciute, D., Mackeviciene, I., Kuslys, M., & Correia, R. F. (2023). The role of influencer–follower congruence in the relationship between influencer marketing and purchase behaviour. *Journal of Retailing and Consumer Services*, 75, 103506. <https://doi.org/10.1016/j.jretconser.2023.103506>
- Vrontis, D., Makrides, A., Christofi, M., & Thrassou, A. (2021). Social media influencer marketing: A systematic review, integrative framework and future research agenda. *International Journal of Consumer Studies*, 45(4), 617–644. <https://doi.org/10.1111/ijcs.12647>
- Wahid, R., Karjaluoto, H., Taiminen, K., & Asiati, D. I. (2023). Becoming tiktok famous: Strategies for global brands to engage consumers in an emerging market. *Journal of International Marketing*, 31(1), 106–123. <https://doi.org/10.1177/1069031X221129554>

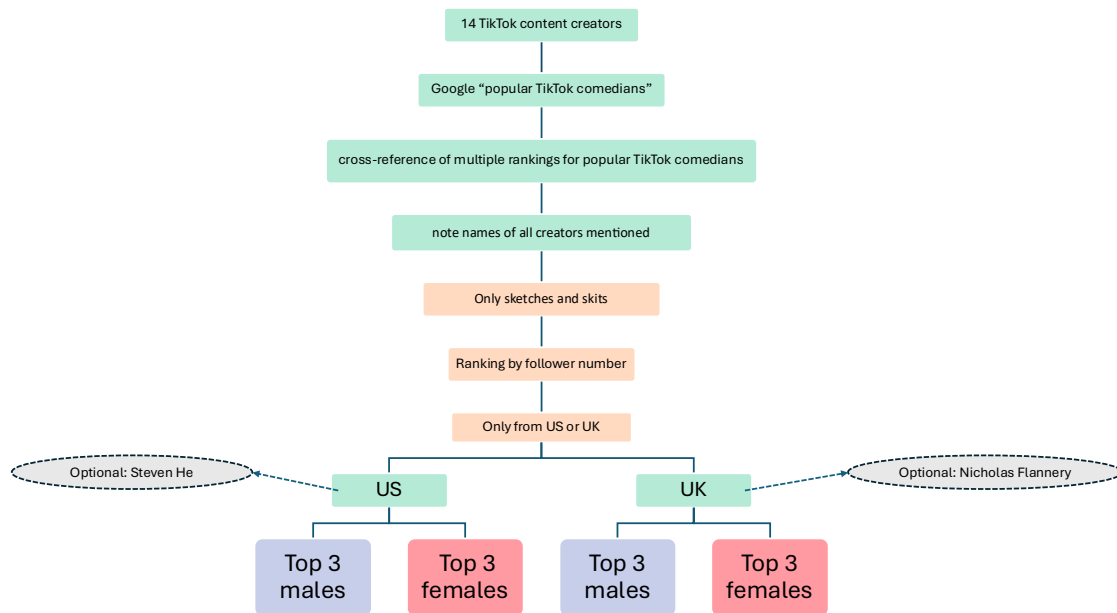
- Wan, Q., & Lu, Z. (2024). Investigating vtubing as a reconstruction of streamer self-presentation: Identity, performance, and gender. *Proceedings of the ACM on Human-Computer Interaction*, 8(CSCW1), 1–22. <https://doi.org/10.1145/3637357>
- Watier, N. (2023). Subtle threat cues in marketing horror and children’s entertainment. *Psychology of Popular Media*, 12(2), 231–241. <https://doi.org/10.1037/ppm0000401>
- Wen, X., Wang, M., Richardt, C., Chen, Z.-Y., & Hu, S.-M. (2020). Photorealistic audio-driven video portraits. *IEEE Transactions on Visualization and Computer Graphics*, 26(12), 3457–3466. <https://doi.org/10.1109/TVCG.2020.3023573>
- Wickham, H. (2009). *Stringr: Simple, consistent wrappers for common string operations* [Data set]. <https://doi.org/10.32614/CRAN.package.stringr>
- Wood, S. (2000). *Mgcv: Mixed gam computation vehicle with automatic smoothness estimation* [Data set]. <https://doi.org/10.32614/CRAN.package.mgcv>
- Wood, S., & Scheipl, F. (2009). *Gamm4: Generalized additive mixed models using ‘mgcv’ and ‘lme4’* [Data set]. <https://doi.org/10.32614/CRAN.package.gamm4>
- Wu, T., Reynolds, J., Schlegelmilch, B. B., & Wu, J. (2025). How ceos can increase public engagement via microblogs: The role of ceo gender and content orientation. *British Journal of Management*, 36(1), 130–144. <https://doi.org/10.1111/1467-8551.12830>
- Xiao, L., Li, X., & Mou, J. (2024). Exploring user engagement behavior with short-form video advertising on short-form video platforms: A visual-audio perspective. *Internet Research*. <https://doi.org/10.1108/INTR-07-2023-0521>
- Xiao, L., Li, X., & Zhang, Y. (2023). Exploring the factors influencing consumer engagement behavior regarding short-form video advertising: A big data perspective. *Journal of Retailing and Consumer Services*, 70, 103170. <https://doi.org/10.1016/j.jretconser.2022.103170>
- Xiao, L., Sun, L., Lei, Z., & Fu, B. (2025). How do features in health short-form videos impact viewer engagement: An empirical study. *Industrial Management & Data Systems*, 1–32. <https://doi.org/10.1108/IMDS-04-2024-0385>

- Xie, Z., Korrapolu, N. K., Dubey, A., Song, L., Xu, C., Wilson, K. M., Cupertino, A., & Li, D. (2025). Leveraging large language models to identify engagement-driving features in vaping-related tiktok videos: Cross-sectional study. *Journal of Medical Internet Research*, 27, e76265–e76265. <https://doi.org/10.2196/76265>
- Xing, J. (2024). A multidimensional perspective on cross-gender videos in tiktok: Performance strategies, content production, and social impacts. In Z. Zhan, J. Liu, D. M. Elshenawi, & E. Duester (Eds), *Proceedings of the 2024 10th International Conference on Humanities and Social Science Research (ICHSSR 2024)* (Vol. 858, pp. 1729–1735). Atlantis Press SARL. https://doi.org/10.2991/978-2-38476-277-4_194
- Xu, Q., Kim, E., & Guo, S. (2024). Digital selves: A cross-cultural examination of athlete social media self-presentation during the 2020 Tokyo Olympics. *International Review for the Sociology of Sport*, 59(5), 726–746. <https://doi.org/10.1177/10126902231224438>
- Xu, T., Liu, Y., Jin, Y., Qu, Y., Bai, J., Zhang, W., & Zhou, Y. (2025). From recorded to AI-generated instructional videos: A comparison of learning performance and experience. *British Journal of Educational Technology*, 56(4), 1463–1487. <https://doi.org/10.1111/bjet.13530>
- Yang, Q., Wang, Y., Wang, Q., Jiang, Y., & Li, J. (2025). Harmonizing sight and sound: The impact of auditory emotional arousal, visual variation, and their congruence on consumer engagement in short video marketing. *Journal of Theoretical and Applied Electronic Commerce Research*, 20(2), 69. <https://doi.org/10.3390/jtaer20020069>
- Yin, X., Li, J., Si, H., & Wu, P. (2024). Attention marketing in fragmented entertainment: How advertising embedding influences purchase decision in short-form video apps. *Journal of Retailing and Consumer Services*, 76, 103572. <https://doi.org/10.1016/j.jretconser.2023.103572>
- Yu, Y., Chen, H., Romero, D., & Dhillon, P. S. (2025). Signals in the noise: Decoding unexpected engagement patterns on twitter. *Proceedings of the ACM on Human-Computer Interaction*, 9(7), 1–33. <https://doi.org/10.1145/3757657>

- Zannettou, S., Nemes-Nemeth, O., Ayalon, O., Goetzen, A., Gummadi, K.P., Redmiles, E.M. and Roesner, F. 2024. Analyzing User Engagement with TikTok's Short Format Video Recommendations using Data Donations *In: Proceedings of the CHI Conference on Human Factors in Computing Systems* (Online). Honolulu HI USA: ACM, pp.1–16. (Accessed 31 July 2025). Available from: <https://dl.acm.org/doi/10.1145/3613904.3642433>.
- Zhang, C., Zheng, H., & Wang, Q. (2022). Driving factors and moderating effects behind citizen engagement with mobile short-form videos. *IEEE Access*, 10, 40999–41009. <https://doi.org/10.1109/ACCESS.2022.3167687>
- Zhang, J., & Li, C. (2025). The influence and configuration effect of content characteristics on customer input in the context of short video platforms. *Journal of Research in Interactive Marketing*, 19(3), 482–497. <https://doi.org/10.1108/JRIM-02-2024-0121>
- Zhang, J., Synave, R., Delepoulle, S., & Cozot, R. (2023). Reconstructing image composition: Computation of leading lines. *Journal of Imaging*, 10(1), 5. <https://doi.org/10.3390/jimaging10010005>
- Zhang, X., Zhao, Z., & Wang, K. (2024). The effects of live comments and advertisements on social media engagement: Application to short-form online video. *Journal of Research in Interactive Marketing*, 18(3), 485–505. <https://doi.org/10.1108/JRIM-02-2023-0069>
- Zhao, H., & Wagner, C. (2023). How TikTok leads users to flow experience: Investigating the effects of technology affordances with user experience level and video length as moderators. *Internet Research*, 33(2), 820–849. <https://doi.org/10.1108/INTR-08-2021-0595>
- Zhao, Y., Qiao, T., Chen, Y., Kuang, M., Bai, W., Yi, Y., Huang, X., Li, W., & Wang, W. (2025). Attention on social media depends more on how you express yourself than on who you are. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-025-02323-1>

Appendices (Chapter 3):

1. Sampling Procedures



2. Supplementary tables: Gender-Differentiated Effects of Role-Play on Production Elements

Table 15 Gender by Role-Play Type Interactions for Production Elements (continuous variables)

Production Element	Role-Play Type	Female Creators (Reference)	Male Creators	Gender Interaction		
		β (SE)	β (SE)	β (SE)	p	
Linguistic/Textual						
Duration (seconds)	Same-gender	29.43 (4.65)***	3.41 (5.70)	26.02 (7.36)	< .001	***
	Cross-gender	11.21 (5.38)*	-7.22 (9.59)	18.43 (10.99)	0.094	†
Word Count (Title)	Same-gender	3.87 (1.21)**	-0.02 (1.45)	-3.89 (1.55)	0.013	*
	Cross-gender	7.39 (1.56)***	6.10 (2.25)	** -1.29 (2.31)	0.578	
Title Sentiment	Same-gender	0.53 (0.12)***	0.04 (0.14)	-0.48 (0.17)	0.004	**
	Cross-gender	0.40 (0.18)*	0.47 (0.24)	* 0.07 (0.25)	0.780	
First-Person Pronoun Ratio	Same-gender	-0.010 (0.003)**	0.004 (0.004)	-0.015 (0.005)	0.006	**
	Cross-gender	-0.007 (0.005)	0.006 (0.009)	-0.012 (0.010)	0.233	
Audio						
Zero Crossing Rate	Same-gender	-0.013 (0.003)***	-0.032 (0.005)	*** -0.018 (0.004)	< .001	***
	Cross-gender	-0.012 (0.005)*	-0.030 (0.008)	*** -0.018 (0.007)	0.006	**
Mean Pitch (semitones)	Same-gender	0.20 (0.29)	1.03 (0.35)	** 0.83 (0.40)	0.037	*
	Cross-gender	0.78 (0.44)	1.63 (0.59)	** 0.85 (0.59)	0.151	
SD Pitch (semitones)	Same-gender	0.18 (0.13)	0.52 (0.18)	** 0.34 (0.17)	0.049	*
	Cross-gender	0.24 (0.19)	0.36 (0.27)	0.13 (0.26)	0.629	
RMS Amplitude	Same-gender	-906.69 (249.54)***	-407.27 (308.54)	499.42 (396.82)	0.209	
	Cross-gender	-1178.30 (286.07)***	-503.23 (513.63)	675.07 (587.92)	0.251	
Visual						
Contrast	Same-gender	0.009 (0.006)	0.007 (0.007)	-0.002 (0.008)	0.775	
	Cross-gender	0.010 (0.006)	-0.014 (0.010)	0.023 (0.012)	0.047	*
Brightness	Same-gender	0.013 (0.019)	-0.005 (0.023)	-0.018 (0.025)	0.468	
	Cross-gender	0.065 (0.035)	0.015 (0.048)	-0.049 (0.040)	0.223	

Table 16 Gender by Role-Play Type Interactions for Production Elements (binary/categorical variables)

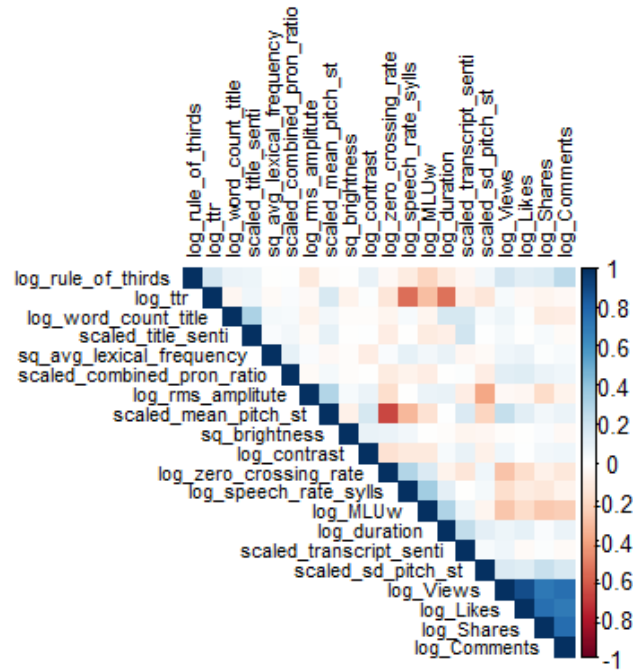
Production Element	Role-Play Type	Female Creators (Reference)		Male Creators		Gender Interaction		
		OR	(95% CI)	OR	(95% CI)	OR	p	
Visual/Technical								
Filter	Same-gender	22.7	(5.49–93.90)***	622.15	(79.87–4846.60)***	27.41	0.011	*
	Cross-gender	8.32	(1.95–35.43)**	368.47	(4.74–28,648.80)**	44.28	0.773	
Green Screen	Same-gender	20.18	(2.43–167.80)**	317.31	(21.35–4716.40)***	15.73	0.086	†
	Cross-gender	28.54	(4.47–182.30)***	24,001,800	(CI not estimable)***	841,163,700	0.743	
Audio								
Background Music	Same-gender	11.49	(2.82–46.78)***	165.02	(28.91–942.00)***	14.37	0.003	**
	Cross-gender	7.12	(0.98–51.58)†	42.14	(2.14–828.20)*	5.92	0.331	
Costume/Styling								
Accessories	Same-gender	90.46	(20.09–407.40)***	27.76	(3.23–238.70)**	0.31	0.329	
	Cross-gender	91.54	(12.44–673.60)***	13.55	(0.85–216.00)	0.15	0.166	
Clothing	Same-gender	4.21	(1.66–10.70)**	4.97	(1.48–16.70)**	1.18	0.845	
	Cross-gender	36.33	(7.47–176.60)***	63.82	(6.13–664.80)***	1.76	0.646	
Hairstyling	Same-gender	3.66	(0.55–24.50)	340.89	(3.03–38,353.10)**	93.12	0.846	
	Cross-gender	21.8	(2.63–180.70)**	20,946,000	(CI not estimable)***	961,000	0.827	
Physical Settings	Same-gender	3.05	(1.31–7.12)**	1.82	(0.57–5.81)	0.6	0.510	
	Cross-gender	0.3	(0.03–2.66)	0.17	(0.01–2.69)	0.57	0.751	
Narrative								
Close Ties	Same-gender	7.35	(2.86–18.88)***	4.21	(1.33–13.36)*	0.57	0.426	
	Cross-gender	26.6	(6.43–110.00)***	36.32	(3.99–330.20)***	1.37	0.734	

Notes

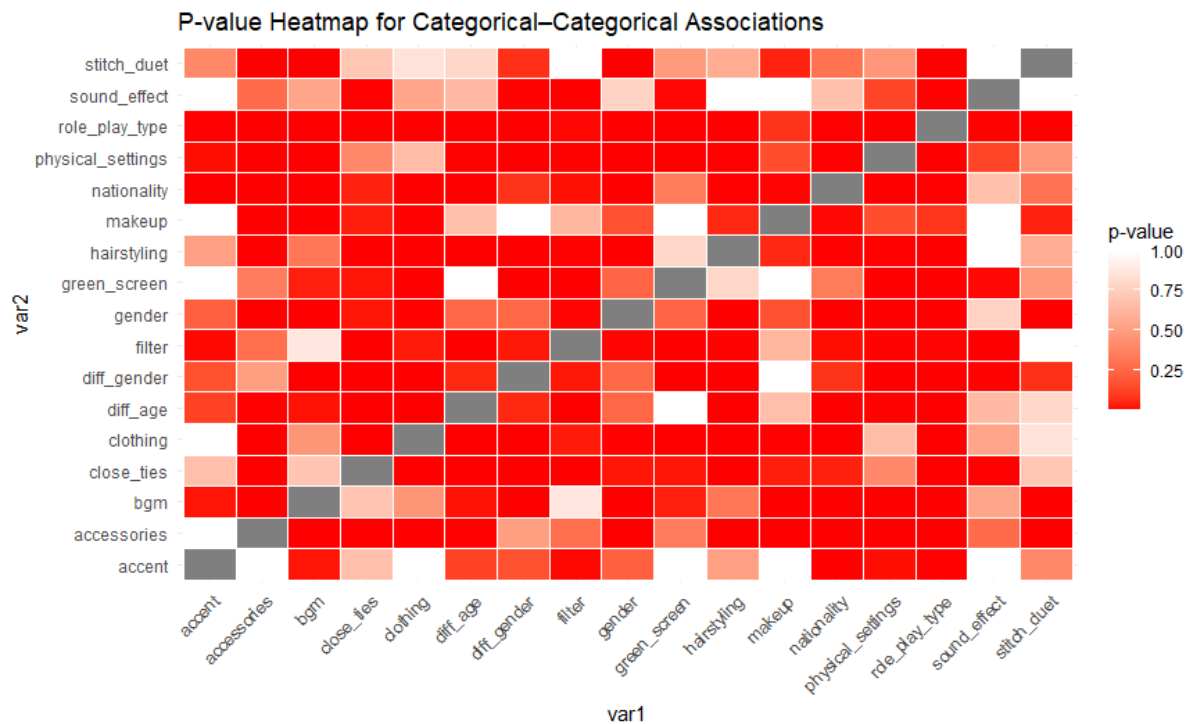
- 1) Female creators serve as the reference category. Interaction terms indicate the multiplicative change in effect for male creators relative to female creators.
- 2) For continuous variables: Positive interaction coefficients indicate that male creators show a larger positive effect (or smaller negative effect) than female creators. Negative coefficients indicate that male creators show a smaller positive effect (or larger negative effect).
- 3) For binary variables: Interaction ORs > 1 indicate that male creators show stronger associations than female creators. ORs < 1 indicate weaker associations for male creators.
- 4) Significance levels: † $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$
- 5) CI = Confidence Interval; OR = Odds Ratio; SE = Standard Error
- 6) Some OR estimates for binary variables exceed typical ranges due to sparse data in specific cells, particularly for green screen and hairstyling in cross-gender role-play. Where noted as "CI not estimable," confidence intervals could not be reliably calculated due to extremely sparse data (zero or near-zero cell counts), resulting in numerical instability. These estimates should be interpreted with caution and are primarily reported for completeness.

3. Correlation between continuous variables

Correlation Matrix of Continous Variables



4. Association between categorical variables



Conclusion

1. Overview

This thesis investigated how multimodal features, production elements, and platform infrastructure jointly shape audience engagement in TikTok comedy content produced during the COVID-19 period (2020–2022), drawing on a shared corpus of videos from 14 UK and US creators. Three interrelated empirical studies were designed to address this aim from complementary angles: the first evaluating the data infrastructure through which platform content is accessed and studied; the second modelling how visual, auditory, linguistic, and platform features jointly predict audience engagement; and the third examining how a specific creative strategy, role-play, independently predicts engagement above and beyond correlated production characteristics. Together, they constitute a progressive analytical sequence in which each paper provides the methodological and conceptual foundation for the next. The present chapter synthesises the principal findings of each paper, evaluates the extent to which the overarching research questions have been addressed, draws out the theoretical and methodological contributions, and identifies directions for future research.

2. Summary of Findings Across the Three Papers

2.1 Paper 1: Platform Data Infrastructure and Its Limitations

Paper 1 addressed a question that is logically prior to any substantive engagement analysis: whether the platform's own Research API can be trusted as a data source for social science research. This is a concern shared across the field. Davidson et al. (2023) have demonstrated that platform-controlled APIs pose systemic risks to open science, and Freelon (2018) identified the structural vulnerabilities of computational research in the post-API age. Paper 1 extended these critiques by subjecting the TikTok Research API to systematic comparative scrutiny, contrasting its outputs against a commercial third-party analytics service (Analisa.io) and manual data collection across a corpus of 6,373 videos from 14 creators (Chapter 2).

The findings confirmed that the API offers genuine strengths. It retrieved a greater volume of videos than either alternative method and provided access to 22 variables, including eight unavailable through any other

source, among them transcriptions, effect identifiers, and precise temporal metadata. These capabilities are consistent with what Corso et al. (2024) identify as the principal advantage of official platform data access: scalability and ethical compliance that scraping-based approaches cannot match.

The limitations, however, were substantial and systematic. Transcription coverage reached approximately ten per cent of content overall, with coverage patterned by nationality and gender in ways that systematically disadvantaged British female creators. Engagement metrics exhibited inconsistent accuracy across data sources. Small changes in sample composition produced disproportionately large shifts in the resulting dataset, indicating a sensitivity to inclusion criteria that researchers must actively manage. These findings extend the critical literature on API bias (Davidson et al., 2023; Freelon, 2018) by providing a worked, quantified audit of precisely what goes wrong when platform data is not triangulated. The paper advocates an integrated multi-method auditing protocol as a necessary corrective, and that protocol is precisely what produced the 741-video corpus on which Papers 2 and 3 depend.

2.2 Paper 2: Multimodal Predictors of Engagement

Paper 2 built on the validated corpus to investigate how features spanning visual composition, audio context, linguistic markers, and platform affordances jointly predict three distinct engagement outcomes. Prior research has predominantly examined these modalities in isolation, a fragmentation that Wang (2021) and Chen (2024) have both identified as a significant limitation of the field. Paper 2 addressed this gap by modelling all feature categories simultaneously using Generalised Additive Mixed Models (GAMMs) with creator-specific random intercepts fitted to 741 comedy videos, with separate models estimated for like count, comment count, and share count.

The central finding was that engagement is irreducibly multidimensional. Likes, comments, and shares are not interchangeable proxies for a single underlying construct. Rather, they represent qualitatively distinct audience behaviours driven by different multimodal configurations, a finding consistent with calls in the literature to move beyond engagement-as-metric towards engagement-as-process (Montag et al., 2021; Gu et al., 2022). Likes were most reliably predicted by visual composition quality and video description length.

Comments were driven primarily by video duration, title sentiment, and linguistic complexity. Shares were shaped by relationship-focused content, transcript sentiment, and lexical properties.

Two findings cut across all three models and merit particular emphasis in this synthesis. Video description length emerged as the sole universal predictor, though with different functional forms across outcomes. This indicates that platform affordances relating to textual metadata operate independently of, and in interaction with, the multimodal content of the video itself, consistent with Bucher and Helmond's (2018) account of how platform features actively shape content reception. Creator gender consistently moderated multimodal effects throughout: female creators' engagement was shaped more by prosodic and linguistic complexity, whilst male creators' engagement depended more on lexical accessibility. This finding extends research by Cheng and Li (2024) and Chung et al. (2023) on gender-differentiated engagement patterns by demonstrating that gender functions as a structural moderating variable rather than a demographic control to be held constant.

2.3 Paper 3: Role-Play, Production Elements, and Gender Effects

Paper 3 narrowed the analytical focus to role-play, a comedic strategy in which creators explicitly embody characters distinct from their everyday personas and the most prevalent creative strategy in the present corpus. Despite its prevalence, prior scholarship on the engagement effects of role-play on short-form video platforms had been largely restricted to qualitative and descriptive accounts (Cervi & Divon, 2023; Divon & Eriksson Krutrök, 2024; Hautea et al., 2021), with limited quantitative evidence on which production elements characterise such content or how it independently predicts audience response. Paper 3 addressed this gap by investigating which production elements systematically distinguish role-play from non-role-play content, and whether role-play independently predicts engagement above and beyond those correlated production characteristics.

The production element analysis, conducted using mixed-effects models with creator-specific random intercepts, revealed that role-play videos employ a recognisable and consistent production signature. They are substantially more likely to feature accessories, distinct costuming, and relationship-focused narratives than non-role-play content, with cross-gender role-play demonstrating even stronger differentiation than

same-gender role-play across all production dimensions. These patterns persist after controlling for creator-level variation, confirming that production elements function as meaningful markers of character embodiment rather than incidental stylistic choices. This is consistent with Gran's (2025) account of how costumes, props, and symbols operate as theatrical signalling devices in social media self-presentation, enabling audiences to recognise and parse the performative mode being employed.

The engagement analysis established that role-play exerts independent effects on likes and shares even after statistically controlling for all associated production elements, demonstrating that the performance strategy itself, rather than its correlated production features, drives audience response. Critically, role-play showed no significant independent effect on comment counts. This differential pattern replicates and extends the multidimensional engagement finding from Paper 2: different engagement behaviours respond to different creative inputs, and a strategy highly effective at generating passive approval and social amplification need not stimulate discursive participation. The gender interaction in the shares model, in which male creators derive substantially greater sharing gains from same-gender role-play whilst female creators show minimal role-play effects on sharing, extended the gender moderation finding from Paper 2 into a new analytical context. This pattern is consistent with research on gender-incongruent content in digital media (Wu et al., 2025) and with Butler's (1990) account of how departures from normative gender citation attract heightened social attention.

2.4 The Relationship Between the Three Papers

The three papers are nested rather than parallel, and this distinction matters for how their contributions should be understood collectively. Paper 1 establishes the epistemic conditions under which the data used in Papers 2 and 3 can be trusted. The multi-method calibration process advocated in Paper 1 is precisely the process that produced the 741-video corpus; the methodological argument and the substantive analyses are inseparable. Paper 2 then establishes the general multimodal architecture of engagement across that corpus, identifying which feature configurations predict each type of audience response and demonstrating that gender consistently moderates those relationships. Paper 3 uses that architecture as a controlled baseline, enabling a much stronger claim than would otherwise be possible: that role-play drives engagement in its own right, independently of the production features with which it is correlated. The logic of the sequence is progressive; each paper depends on, and builds upon, the work that precedes it.

3. Addressing the Research Questions

The three overarching research questions stated in Section 4 of the Introduction can now be evaluated against the evidence gathered. Each question corresponds to one of the three papers in this thesis and is addressed in turn below. Examiners wishing to cross-reference the original question statements should refer to Sections 4.1 through 4.3 of the Introduction.

3.1 RQ1: Methodological Infrastructure and API Validity

The first research question, stated in full in Introduction Section 4.1, was as follows.

RQ1: What are the primary challenges and opportunities of using TikTok’s Research API for analysing societal trends during the COVID-19 pandemic, and how do its methodological constraints affect data validity?

Paper 1 addresses this question directly and with considerable specificity (Bai & Gu, 2026). At the level of opportunities, the API’s scalability and access to unique variables represent genuine advantages for computational social science. These are particularly valuable for researchers requiring longitudinal datasets covering multiple creators over extended time periods.

At the level of challenges, the evidence is unambiguous. Engagement metric inconsistencies across data sources, demographically patterned transcription coverage gaps, and sample composition sensitivity all represent substantive threats to validity. These challenges are not unique to TikTok. They reflect broader structural concerns about platform-controlled data access documented across the field (Davidson et al., 2023; Freelon, 2018; Corso et al., 2024). What Paper 1 adds is a quantified demonstration of their magnitude and a replicable auditing methodology designed to mitigate them.

The multi-method approach adopted here substantially reduced the impact of these limitations. Still, some residual uncertainty about data quality inevitably remains, and the findings of Papers 2 and 3 should be interpreted with that caveat in mind.

3.2 RQ2: Multimodal Predictors of Engagement

The second research question, stated in full in Introduction Section 4.2, was as follows.

RQ2: How do visual composition, audio context, linguistic markers, platform features, and creator characteristics jointly predict engagement metrics in TikTok comedic videos?

Paper 2 addresses this question through a comprehensive multimodal analysis using Generalised Additive Mixed Models (GAMMs) with creator-specific random intercepts, separately modelling like count, comment count, and share count. The evidence gathered is clear and robust within the corpus studied. Multimodal features jointly predict engagement, and they do so in ways that differ systematically across the three outcome metrics.

This finding directly addresses the fragmentation critique of prior unimodal research (Chen, 2024; Wang, 2021; Xiao et al., 2025). It demonstrates that modalities interact rather than simply add, and that their effects are non-linear and context-dependent. The GAMM framework was well-suited to capturing these relationships, accommodating non-linearity without pre-specifying functional forms and addressing the hierarchical structure of videos nested within creators.

Where the findings are less certain is in their generalisability. The corpus of 14 high-profile creators cannot be assumed to represent the full range of TikTok comedy content. The observational design precludes definitive causal inference. Effect estimates should therefore be treated as descriptive findings within this corpus rather than universal parameters.

3.3 RQ3: Production Elements and Role-Play Performance

The third research question, stated in full in Introduction Section 4.3, was as follows.

RQ3: How do production elements and role-play features jointly predict engagement in TikTok comedic videos, and do these effects vary by creator gender?

Paper 3 addresses this question through a two-stage analytical design. The evidence gathered is statistically robust and theoretically interpretable. Role-play content is characterised by a consistent production signature that extends well beyond what individual creator preferences would predict. This confirms that character embodiment constitutes a coherent creative practice with a systematic material logic.

Role-play independently predicts substantially higher likes and shares after controlling for the full production feature set. Creative strategy, rather than production investment per se, drives these outcomes.

The null finding for comments replicates and enriches the multidimensional engagement framework from Paper 2. The gender interaction in the shares model, in which male creators benefit disproportionately from same-gender role-play, is both statistically significant and theoretically interpretable. It aligns with Butler's (1988) account of gender performativity and Wu et al.'s (2025) findings on gender-incongruent content in digital contexts.

The key limitation remains generalisability. Role-play's prevalence in this corpus may partly reflect the specific temporal and cultural context of the pandemic period. Whether the engagement effects identified here hold across different platform contexts or audience demographics remains an open question that future research should address.

4. The UK/US Comparative Dimension

The thesis was designed with an explicit transatlantic comparative dimension, and it is important to address candidly what that design yielded. At the infrastructural level, the UK/US contrast was genuinely informative. Paper 1 identified nationally patterned biases in API coverage: American male creators were systematically over-represented in transcription availability, whilst British female creators received the lowest transcription coverage of any demographic subgroup in the sample (Bai & Gu, 2026). This finding has direct implications for any study relying on TikTok Research API transcription data to draw conclusions about British creators or female content producers more broadly.

At the level of engagement patterns and production practices, the comparative findings were more limited. With seven American and seven British creators in a 14-creator sample, the statistical power available for detecting nationality effects in Papers 2 and 3 was modest, and no robust nationally differentiated engagement patterns emerged that could be cleanly separated from individual creator variation. This is a sample size constraint rather than a conceptual failure. A more adequately powered comparative study, with a substantially larger and better-matched sample, would be needed to determine whether the engagement architecture identified here varies systematically across national contexts, a direction explicitly identified in the future research agenda below.

5. Theoretical Contributions

5.1 Multimodal Communication Theory

Papers 2 and 3 collectively contribute to multimodal communication theory (Kress & Van Leeuwen, 2020).

They demonstrate that engagement in short-form video emerges from coordinated, non-linear interactions among communicative channels rather than from the additive effects of individual modalities. This directly addresses the fragmentation critique of prior unimodal research (Chen, 2024; Wang, 2021). Individual modalities cannot be understood in isolation; engagement prediction requires analytical frameworks capable of capturing coordinated, non-linear interactions across channels simultaneously.

The differentiated engagement framework developed in Paper 2 represents a further substantive contribution. Likes, comments, and shares respond to qualitatively distinct multimodal configurations rather than to a single underlying engagement factor. This extends and refines prior work on the multidimensional character of social media engagement (De Oliveira Santini et al., 2020; Montag et al., 2021). What the present thesis adds is the identification of specific multimodal configurations that predict each engagement dimension, not merely the assertion that dimensions differ.

The finding that modalities interact asymmetrically rather than additively is theoretically significant in its own right. It reinforces the case for analytical frameworks capable of detecting non-linear interaction effects across multiple communicative channels. Engagement research that examines features in isolation will systematically misestimate their contribution and miss the interdependencies that this thesis identifies.

5.2 Performance Theory and Gender Performativity

Paper 3 makes a specific and empirically grounded theoretical argument about Goffman's (1990) dramaturgical framework. Goffman's account of self-presentation as performance is well-established, and its application to social media has been developed productively by Marwick and boyd (2011) and others. Yet Goffman's framework does not straightforwardly accommodate explicit character performance. In TikTok role-play, the gap between performer and persona is itself foregrounded as a source of comedic meaning. This is not impression management in Goffman's sense. It is a performative mode that depends on the audience recognising that everyday self-presentation is being deliberately transgressed.

The evidence for this theoretical extension is empirical rather than merely conceptual. Role-play's significant independent engagement effects, demonstrated after controlling for the full production feature set, confirm that character embodiment generates audience response that cannot be attributed to production investment or multimodal quality alone. This is consistent with Gran's (2025) account of character performance as a distinct communicative act. It extends existing performance theory into short-form digital contexts in a way that is both theoretically motivated and empirically supported.

The gender-atypicality interpretation of the male creator sharing effect connects the findings to Butler's (1988) account of gender performativity. Character embodiment functions, at least for male creators in this corpus, as a socially licensed space for gender-incongruent behaviour, with measurable consequences for social amplification. This is consistent with Wu et al.'s (2025) research on gender-incongruent content in digital contexts. It also aligns with broader scholarship on how algorithmic environments amplify the social attention that norm-violating content attracts (Boeker & Urman, 2022; Guinaudeau et al., 2022). Together, these findings contribute to an emerging body of work on how gender norms operate within, and are mediated by, platform affordances and recommendation algorithms.

5.3 Platform Data Infrastructure and Methodological Theory

The most methodologically novel theoretical contribution of this thesis concerns the status of research APIs as data sources. APIs are not neutral instruments of data access. They are sociotechnical artefacts that actively shape which cultural practices and creators become visible to outside observers. This argument extends Davidson et al.'s (2023) critique of platform-controlled APIs and Freelon's (2018) analysis of post-API computational research. Paper 1 provides a worked, quantified demonstration of how these biases manifest and what they cost in terms of research validity.

The contribution is not solely critical. By demonstrating what systematic multi-method auditing can reveal, Paper 1 makes a constructive methodological case for the integrated auditing protocol it proposes. In this respect, the thesis contributes to an emerging literature on platform research design (Corso et al., 2024; Entrena-Serrano et al., 2025), providing a replicable framework that future studies can adopt and refine. The theoretical and the methodological are, in platform research, inseparable. Understanding what a platform

does requires understanding both the creative practices it enables and the infrastructural conditions under which those practices become observable to researchers.

6. Methodological Reflections

The methodological contribution of this thesis is concentrated in Paper 1, but its implications reverberate across all three studies. The three-source triangulation design, combining API retrieval, third-party analytics via Analisa.io, and systematic manual verification, detected inconsistencies that no single-source approach could have identified. These included systematic engagement metric divergences across data sources, nationally patterned transcription coverage gaps, and pronounced sensitivity to sample composition. These are concrete, quantified findings that establish what is at stake when researchers rely on platform APIs without independent verification, consistent with calls in the literature for more rigorous multi-method approaches to platform data collection (Davidson et al., 2023; Bai & Gu, 2026).

The GAMM framework deployed in Papers 2 and 3 is itself a methodological contribution. It handles three features of this dataset that prior work has typically neglected: non-linear relationships between continuous predictors and engagement outcomes, hierarchical data structure arising from the nesting of videos within creators, and gender moderation of multimodal effects. The framework's flexibility, achieved through penalised spline smoothing, enables detection of curvilinear patterns without requiring their functional form to be specified in advance, and it does so whilst maintaining interpretability. The AIC-guided model selection procedure, combined with variance inflation factor checks and False Discovery Rate correction in Paper 3, provides a level of analytical rigour appropriate to the complexity of the data.

Taken together, the experience of all three papers supports a practical and staged recommendation for researchers wishing to use TikTok data for social science research. TikTok Research API should not be used as a sole data source. Paper 1 demonstrates that API outputs carry systematic demographic biases and engagement metric inconsistencies that single-source designs cannot detect (Bai & Gu, 2026). A three-source design combining API retrieval, a commercial analytics service, and systematic manual verification is advisable for producing a corpus with known and bounded data quality. This is the design that produced the

741-video corpus on which Papers 2 and 3 depend. The reliability of those papers' findings is inseparable from that prior auditing step.

Within this three-source framework, the appropriate weighting across sources varies by the type of analysis a study intends to conduct. For engagement metric modelling of the kind carried out in Chapters 3 and 4, commercial analytics data provide a useful cross-check on API-reported values, and any systematic divergence between the two should be investigated before modelling proceeds. For linguistic and transcript-based analyses, the API's approximately ten per cent transcription coverage represents the primary constraint. Supplementing API transcriptions with purpose-deployed speech-to-text services substantially increases coverage and reduces the demographic skew in available linguistic data, as the multi-method protocol in Chapter 2 demonstrates. For visual and audio feature extraction, the video files retrieved via the API can be analysed computationally with relatively high coverage, since these analyses do not depend on transcription availability. Studies focused primarily on visual or audio features therefore face a less severe data quality problem than studies centred on linguistic features, though engagement metric verification remains advisable across all study types. The critical practical implication is that the choice of analytical focus should be made with explicit awareness of the coverage and reliability profile of each data stream, so that the scope of any resulting claims is calibrated to what the verified data can actually support.

Several limitations should nonetheless be acknowledged plainly. The 14-creator sample, though selected strategically for diversity across gender and nationality, is too small to support strong generalisations across the population of TikTok comedy creators. The approximately ten per cent transcription coverage limits the linguistic analyses in Papers 2 and 3 to content with recoverable verbal dimensions. The observational design precludes definitive causal inference: the engagement effects estimated are associational, and experimental manipulation of specific features would be required to establish causality. The dataset begins on 1 January 2020, and the first months of the corpus, before UK and US lockdowns were imposed in late March 2020, provide a limited within-sample pre-restriction baseline; however, this is not equivalent to a purpose-designed longitudinal study. These limitations define the scope of the findings rather than

undermining them. The integrated auditing protocol, the multimodal engagement architecture, and the role-play production signatures identified in this thesis are robust within the corpus studied.

7. Policy, Ethical, and Platform Governance Implications

The findings carry implications that extend beyond the academic literature and into questions of platform governance, research infrastructure, and digital rights.

For platform governance, Paper 1's findings make a straightforward but important argument: opaque API design has measurable consequences for research validity and demographic representativeness. When API transcription coverage is systematically lower for British female creators, the research literature that relies on that API will systematically under-represent their comedic expression. Platforms that wish to support credible research should be required to document and disclose these biases rather than leaving researchers to discover them through post hoc auditing. The EU Digital Services Act's provisions on data access for vetted researchers represent a meaningful step in this direction, though implementation will require ongoing scrutiny (European Centre for Algorithmic Transparency, 2025).

For recommendation algorithm design, the differential engagement effects identified in Papers 2 and 3 suggest that algorithms calibrated primarily on aggregate engagement metrics will systematically advantage certain multimodal configurations and certain creator demographics. The structural differences in engagement patterns between male and female creators are not individual idiosyncrasies; they are features of the platform's engagement architecture that algorithm optimisation will tend to amplify rather than neutralise. Platform designers should recognise these differential dynamics, particularly where equity of content exposure is a stated platform objective.

Ethically, the use of publicly available creator content at scale raises questions that this thesis has not fully resolved. The 14 creators in this sample produced their content in a public space, and their TikTok handles are used in Paper 1 for transparency. Nevertheless, the detailed characterisation of individual creators' production strategies and engagement patterns in Papers 2 and 3, without their active participation, raises questions about consent and representation that the field has not adequately settled (Entrena-Serrano, 2025).

Future research programmes involving the systematic analysis of named creators should consider whether participatory approaches are feasible and appropriate.

8. Future Research Agenda

Five directions for future research emerge from the findings, limitations, and supervisory recommendations associated with this thesis. Each is motivated by a specific gap or constraint identified above

Direction 1: Extending the Pipeline to Non-Verbal Content

The first direction concerns extending the multimodal analytical pipeline to non-verbal content. The 741-video dataset includes only videos with automated transcriptions, covering approximately ten per cent of the full API-retrieved sample. This limits the framework to content with a recoverable verbal dimension, leaving non-verbal comedic expression outside its current scope. Incorporating computer vision models capable of extracting facial expression features at scale would substantially broaden the evidential base.

Preliminary exploratory work using models available through Hugging Face suggests that a meaningful proportion of comedic effect in the corpus is conveyed through facial expression rather than speech. Facial sentiment signals often diverge from both audio and transcript sentiment in theoretically interpretable ways. This leads directly to a related priority: systematic investigation of how humour is achieved through multimodal incongruity among facial, audio, and linguistic sentiment channels. Early analyses suggest some humour operates through irony, in which all three channels converge, and other humour through sarcasm, in which they diverge. How gender moderates these two incongruity structures, and how they are distributed across role-play content, are questions the existing data and emerging computational tools make tractable.

Direction 2: Comparative and Cross-Platform Studies

The second direction encompasses comparative and cross-platform studies. The Chinese Douyin corpus collected during the original sampling process remains unanalysed. Applying the multimodal pipeline from Papers 1 through 3 to that data would extend the findings to a culturally and politically distinct platform context. It would also enable a direct TikTok/Douyin comparison, testing whether the engagement

architecture and gender moderation effects hold across platforms sharing technical infrastructure but operating under different regulatory and cultural conditions.

A more adequately powered UK/US comparative study would complement this work. The 14-creator sample could not provide the statistical power required to resolve nationality effects with confidence. A sample of at least 30 creators per national context, matched on follower size, posting frequency, and content genre, would address that gap directly.

Direction 3: Broadening the Sample

Third, the scope of the sample should be broadened in three respects. The present thesis focused on a specific creator profile within a single content genre. Future research should examine TikTok comedy across other formats, including sketch comedy, lip-sync parody, and reaction content. Testing whether the multimodal engagement architecture identified in Paper 2 generalises across these formats, or whether distinct formats carry distinct engagement profiles, remains an open question.

Future work should also incorporate a wider range of creator characteristics. Gender and nationality were the primary creator-level variables in Papers 2 and 3. Additional attributes, including age, ethnicity, follower tier, and creator tenure, may moderate the multimodal engagement relationships identified here in ways the present sample was not designed to detect.

A third priority is expanding the geographical and linguistic scope of the research. The present thesis examines creators from the United Kingdom, the United States, and, in preliminary comparative work, China. Extending to other national contexts, including non-English-speaking countries, is a logical and necessary next step. Yet this expansion raises a validity challenge that must be addressed directly. Automated speech-to-text services perform markedly less well on non-English audio. Any expansion of the framework to multilingual corpora would need to evaluate transcription quality explicitly, following the audit logic established in Paper 1. This is not merely a technical constraint. It is a validity concern that bears directly on the reliability of the linguistic features analysed in Papers 2 and 3.

Direction 4: Integrating Creator and Audience Perspectives

Fourth, integrating creator and audience perspectives would substantially enrich the analytical framework. The present thesis analyses creator content and platform-recorded engagement metrics without direct access to audience characteristics or reception processes. Qualitative interviews with creators would illuminate the motivations underlying role-play adoption. They would also help interpret the gender dynamics that the statistical models identify but cannot explain from within the data alone.

Consumer-side survey or panel data would allow researchers to test the sensory, cognitive, and social verdict framework directly with viewers. This would provide a perceptual foundation for what are currently engagement-based inferences, moving the field from observed behaviour to explained experience.

Direction 5: Improving Platform Data Infrastructure

Fifth, improving platform data infrastructure remains a methodological priority. The integrated API auditing protocol developed in Paper 1 (Bai & Gu, 2026) should be extended into a standardised procedure applicable across multiple platform APIs. It should be capable of being implemented prospectively as platforms evolve under regulatory pressure, including from instruments such as the EU Digital Services Act. The present protocol is a foundation rather than a finished product. The case for its wider adoption is, if anything, stronger now than when the data for this thesis were first collected.

9. Closing Reflection

The introduction to this thesis identified three questions that are likely to remain important well beyond the specific temporal context that frames the data. How should platforms be held accountable for the data they make available to researchers? How do multiple communicative channels interact in shaping the reception of short-form video content? And how do creative strategies, enacted within and constrained by algorithmic environments, succeed or fail in generating different kinds of audience response?

This thesis advances each of them with evidence, method, and theoretical argument. On the first: Paper 1 demonstrates that accountability requires active auditing rather than passive trust and proposes a multi-method framework adequate to that task. On the second: Papers 2 and 3 establish that communicative channels interact asymmetrically and non-linearly, that those interactions are moderated by creator gender,

and that different engagement outcomes respond to distinct multimodal configurations. On the third: Paper 3 demonstrates that role-play, a creative strategy requiring no elaborate technical infrastructure, independently drives substantial engagement gains in a gender-differentiated pattern that is both statistically robust and theoretically interpretable.

Taken together, the three papers furnish quantified, replicable, and theoretically situated answers rather than merely gesturing towards them. Some questions remain open, as they must in any empirical programme of this scope. The generalisability of the findings across creator demographics, content genres, and platform contexts awaits investigation in future work. But the core contributions are substantive. The multi-method auditing protocol, the multimodal engagement architecture, and the role-play production signatures identified across this thesis provide a firm foundation from which that future work can proceed.

Humour, as this thesis has sought to demonstrate, is a more analytically tractable and structurally embedded feature of digital platform life than its apparently ephemeral character might suggest. Understanding how it operates, how it is shaped by platform affordances and data infrastructures, and how audiences respond to it in differentiated ways remains a research agenda well worth pursuing.

References:

- Bai, D., & Gu, Y. (2026). Harnessing big data, hindered by bias: Evaluating TikTok Research API for fair and optimal social sciences. *Social Science Computer Review*, 08944393251413277. <https://doi.org/10.1177/08944393251413277>
- Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.
- Cervi, L., & Divon, T. (2023). Playful activism: Memetic performances of palestinian resistance in tiktok #challenges. *Social Media + Society*, 9(1), 20563051231157607. <https://doi.org/10.1177/20563051231157607>
- Chen, X. (2024). Cross-gender performances and audience engagement on Douyin. *Advances in Social Behavior Research*, 12(1), 31–36. <https://doi.org/10.54254/2753-7102/12/2024120>
- Chung, J. (Jae), Ding, Y., & Kalra, A. (2023). I really know you: How influencers can increase audience engagement by referencing their close social ties. *Journal of Consumer Research*, 50(4), 683–703. <https://doi.org/10.1093/jcr/ucad019>
- Davidson, B. I., Wischerath, D., Racek, D., Parry, D. A., Godwin, E., Hinds, J., van der Linden, D., Roscoe, J. F., Ayravainen, L., & Cork, A. G. (2023). Platform-controlled social media APIs threaten open science. *Nature Human Behaviour*, 7(12), 2054–2057. <https://doi.org/10.1038/s41562-023-01750-2>
- Divon, T., & Eriksson Krutrök, M. (2024). Playful trauma: Tiktok creators and the use of the platformed body in times of war. *Social Media + Society*, 10(3), 20563051241269281. <https://doi.org/10.1177/20563051241269281>
- Entrena-Serrano, C. (2025). Watch, scroll, repeat: How interface design shapes consumptive curation affordances on TikTok. *Social Media + Society*, 11(3). <https://doi.org/10.1177/20563051251358529>
- European Centre for Algorithmic Transparency. (2025). *FAQs: DSA data access for researchers*. European Commission. https://algorithmic-transparency.ec.europa.eu/news/faqs-dsa-data-access-researchers-2025-07-03_en
- Fogarty, M., & Elliot, D. L. (2020). The role of humour in the social care professions: An exploratory study. *The British Journal of Social Work*, 50(3), 778–796. <https://doi.org/10.1093/bjsw/bcz027>

- Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, 35(4), 665–668.
<https://doi.org/10.1080/10584609.2018.1477506>
- Goffman, E. (1990). *The presentation of self in everyday life*. Penguin.
- Gran, A.-B. (2025). Performing not-not-me in some: A new theatrical typology of self-presentation online. *Social Media + Society*, 11(1), 20563051251315256. <https://doi.org/10.1177/20563051251315256>
- Guinaudeau, B., Munger, K., & Votta, F. (2022). Fifteen Seconds of Fame: TikTok and the Supply Side of Social Video. *Computational Communication Research*, 4(2), 463–485.
<https://doi.org/10.5117/CCR2022.2.004.GUIN>
- Hautea, S., Parks, P., Takahashi, B., & Zeng, J. (2021). Showing they care (Or Don't): Affective publics and ambivalent climate activism on tiktok. *Social Media + Society*, 7(2), 20563051211012344.
<https://doi.org/10.1177/20563051211012344>
- Kress, G., & Van Leeuwen, T. (2020). *Reading images: The grammar of visual design* (3rd edn). Routledge.
<https://doi.org/10.4324/9781003099857>
- Montag, C., Yang, H., & Elhai, J. D. (2021). On the psychology of tiktok use: A first glimpse from empirical findings. *Frontiers in Public Health*, 9, 641673. <https://doi.org/10.3389/fpubh.2021.641673>
- Shifman, L. (2015). *Memes in digital culture*. MIT Press.
- Wang, S., Yang, D., Shehata, B., & Li, M. (2023). Exploring effects of intelligent recommendation, interactivity, and playfulness on learning engagement: An application of TikTok considering the meditation of anxiety and moderation of virtual reward. *Computers in Human Behavior*, 149, 107951. <https://doi.org/10.1016/j.chb.2023.107951>
- Wang, S., Yang, D., Shehata, B., & Li, M. (2023). Exploring effects of intelligent recommendation, interactivity, and playfulness on learning engagement: An application of TikTok considering the meditation of anxiety and moderation of virtual reward. *Computers in Human Behavior*, 149, 107951. <https://doi.org/10.1016/j.chb.2023.107951>

- Wang, Y. (2020). Humor and camera view on mobile short-form video apps influence user experience and technology-adoption intent, an example of TikTok (DouYin). *Computers in Human Behavior*, 110, 106373. <https://doi.org/10.1016/j.chb.2020.106373>
- Wu, T., Reynolds, J., Schlegelmilch, B. B., & Wu, J. (2025). How ceos can increase public engagement via microblogs: The role of ceo gender and content orientation. *British Journal of Management*, 36(1), 130–144. <https://doi.org/10.1111/1467-8551.12830>
- Zekavat, M. (2025). Laughter in times of distress: Pandemic humor and satire in Iran. *European Journal of Cultural Studies*, 28(6), 1799–1823. <https://doi.org/10.1177/13675494251331211>