



**UNIVERSITY
OF TURKU**

Turku School of
Economics

Multimodal Short Video Strategies of Independent Artists Navigating Algorithmic Gatekeeping on TikTok

Master's Thesis in Marketing

Author:
Noora Suurhasko

Supervisor:
D.Sc Harri Terho

10.4.2026
Turku

Student's statement regarding the use of Artificial Intelligence (AI) for preparing and/or writing this thesis:

I have not used any AI-based tools.

I have used AI-based tools. Their use is documented in the Appendix. The AI tools were used in a way that complies with academic integrity guidelines.

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service.

Master's thesis

Subject: Marketing

Author: Noora Suurhasko

Title: Multimodal Short Video Strategies of Independent Artists Navigating Algorithmic Gatekeeping on TikTok

Supervisor: D.Sc Harri Terho

Number of pages: 87 pages (+ appendices 21 pages)

Date: 10.4.2026

Abstract

Independent artists today navigate a digital music market defined by oversupply and algorithmic gatekeeping—conditions in which visibility is structurally scarce, not incidentally so. On TikTok, short-form videos must achieve immediate interpretability to circulate within an interest-graph feed, yet the design logic behind high-engagement TikTok posts has received limited scholarly attention. Existing research has examined platform dynamics, branding, and short-form communication in relative isolation. This study examines how independent artists design multimodal content within platform-native short-form video posts to generate engagement, advancing an artist-branding perspective on content design as a sociomaterial practice.

The study is guided by three research questions that move from surface to structure to strategy: what content features and communicative messages characterize high-engagement posts; how those features organize narrative meaning through micro-storytelling sequences; and what strategic archetypes of artist-brand communication emerge when both are considered together. The analysis integrates a sociomaterial perspective on platform visibility with a stimulus–staging–response (S–St–R) framework, distinguishing among stimulus (multimodal surface cues), staging (their temporal sequencing), and response (platform-visible audience engagement). In combination, these lenses reframe short-form content design as a relational process at the level of content design—one in which the relations between elements carry analytical weight equal to the elements themselves.

To examine these questions, the study adopts a qualitative, abductive approach, drawing on a purposive sample of 30 high-engagement TikTok videos published by 10 anonymized independent artists in early 2025. Each post is treated as a multimodal unit of analysis and analyzed through Visual-Verbal Video Analysis (VVVA)—a method purpose-built for holding audiovisual cues, micro-temporal narrative structures, and engagement traces within a single analytical frame.

The findings demonstrate that high-engagement TikTok content appears to rely on recurring multimodal patterns that establish immediate legibility. Meaning is structured through three micro-storytelling patterns—Immediate Loops, Revealed Progressions, and Escalation Arcs—which, combined with message-level orientations toward entertainment value and emotional value, yield four analytically distinct strategic archetypes of artist-brand communication: Immediate Musical/Aesthetic Loops, Proof-to-Reveal Configurations, Process-to-CTA Funnels, and Escalation-to-Punchline Formats. The study introduces configuration fluency as the mechanism by which independent artists align multimodal elements, narrative sequencing, and platform-native affordances, proposing legibility-first content design as a practical principle for algorithmically mediated environments and offers a basis for future research at the intersection of platform studies, multimodal communication, and independent artist practice.

Keywords: Independent artists; TikTok; Artist branding; Algorithmic gatekeeping; Multimodal content design; Micro-storytelling; Configuration fluency; Visual–Verbal Video Analysis (VVVA)

Pro gradu -tutkielma

Oppiaine: Markkinointi

Tekijä: Noora Suurhasko

Otsikko: Itsenäisten artistien lyhytvideostrategiat TikTokissa algoritmisen portinvartijuuden aikakaudella

Ohjaaja: KTT Harri Terho

Sivumäärä: 87 sivua (+ liitteet 21 sivua)

Päivämäärä: 10.4.2026

Tiivistelmä

Nykypäivän digitaalisia musiikkimarkkinoita leimaavat sisällön ylitarjoja ja algoritmisen portinvartijuus, jotka tekevät näkyvyydestä rakenteellisesti niukan resurssin itsenäisille artisteille. TikTokin käyttäjäkohtaisesti räätälöity sisältösyöte edellyttää markkinointiviestinnältä välitöntä ymmärrettävyyttä: katsojan on hahmotettava videon idea ja arvo sekunneissa, muuten algoritmi ohittaa sen. Korkea sitoutuminen (katselut, tykkäykset ja jaot) on keskeinen onnistumisen mittari, mutta sen taustalla olevaa sisältölogiikkaa on tutkittu vähän. Aiempi tutkimus on tarkastellut alustojen dynamiikkaa, artistibrändäystä ja lyhytvideoviestintää pääosin erillisinä ilmiöinä.

Tässä tutkielmassa analysoidaan, miten itsenäiset artistit suunnittelevat multimodaalista lyhytvideosisältöä yleisön sitoutumisen vahvistamiseksi. Tutkimus tarkastelee sisällönsuunnittelua sosiomateriaalisena käytäntönä ja kysyy, millaiset piirteet luonnehtivat korkean sitoutumisen julkaisuja, miten ne muodostavat merkityksiä mikrotarinankerronnassa ja millaisia strategisia arkkityyppejä artistibrändiviestinnässä voidaan tunnistaa. Teoreettinen viitekehys yhdistää alustanäkyvyyden sosiomateriaalisen tarkastelun ja ärsyke–jäsentely–vaste-mallin (S–St–R), jossa erotetaan videon multimodaaliset ominaisuudet, niiden ajallinen järjestymisen ja katsojan vaste. Näin lyhytvideosisällön suunnittelu jäsentyy prosessina, jossa elementtien väliset suhteet ovat keskeisiä.

Tutkimus on laadullinen ja abduktiivinen. Aineisto koostuu 30 korkean sitoutumisen TikTok-videosta, jotka on kerätty kymmenen anonymisoidun itsenäisen artistin tileiltä alkuvuodesta 2025 ja analysoitu visuaalis-verbaalisen videoanalyysin (VVVA) avulla. Tulokset osoittavat, että korkean sitoutumisen sisällöt perustuvat toistuviin multimodaalisiin malleihin ja jäsentyvät kolmeen mikrotarinankerronnan jaksoon: toistoon, vaiheittaiseen paljastamiseen ja jännitteen huipentamiseen. Näiden pohjalta voidaan tunnistaa neljä strategista arkkityyppiä: loop-rakenteinen välitön immersio, paljastukseen etenevä kerronta, prosessista toimintakehotukseen johtava suppilo sekä huipennukseen rakentuva dramaturgia.

Tutkielman keskeinen kontribuutio on konfiguraatiosujuvuuden käsite, joka kuvaa, miten multimodaaliset elementit, tarinankerronnan ajallinen rakenne ja alustan tekniset reunaehdot sovitetaan yhteen. Mitä sujuvammin nämä tasot yhdistyvät, sitä tehokkaammin sisältö puhuttelee katsojaa. Tutkielma esittää algoritmiohjattuihin ympäristöihin suunnitteluperiaatteen, joka korostaa sisällön välitöntä ymmärrettävyyttä ja avaa uusia tutkimusmahdollisuuksia alustatutkimuksen, multimodaalisen viestinnän ja itsenäisen artistitoiminnan risteyskohdassa.

Avainsanat: itsenäiset artistit; TikTok; artistibrändäys; algoritmisen portinvartijuus; multimodaalinen sisällönsuunnittelu; mikrotarinankerronta; konfiguraatiosujuvuus; visuaalis-verbaalinen videoanalyysi

TABLE OF CONTENTS

1	Introduction	9
1.1	Relevance of the Study	9
1.2	Research Gap	10
1.3	Purpose of the Study	12
1.4	Delimitations and Structure	12
2	Independent Artist Branding in the Algorithmic Era	14
2.1	Independent Artists under Algorithmic Gatekeeping	14
2.2	Artist Branding on Social Platforms	15
2.3	TikTok's Algorithmic Visibility	17
3	Design of Short-Form Content in Algorithmic Platforms	21
3.1	Engagement as Observable Response Traces	21
3.2	Content Stimulus Building Blocks	22
3.2.1	Content Features	24
3.2.2	Content Messages	26
3.3	Narrative Staging in Short-Form Video	28
3.4	Configuration Logic for Integrating Stimulus, Staging, and Response	31
4	Methodology	35
4.1	Philosophical Approach	35
4.2	Research Design	36
4.3	Data Sample and Unit of Analysis	37
4.4	Visual-Verbal Video Analysis (VVVA)	40
4.5	Operationalization of Analytical Dimensions	40
4.6	Abductive Development of the Coding Framework	44
4.7	Trustworthiness of the Study	45
4.8	Ethical Considerations	47
5	Findings	49
5.1	Content Features and Communicative Messages in High-Engagement TikTok Videos	49

5.1.1	Content Features and Multimodal Form	50
5.1.2	Message-Level Meaning Patterns	52
5.2	Micro-Storytelling Sequences in High-Engagement TikTok Videos	55
5.3	Strategic Archetypes of Artist-Brand Communication in High-Engagement TikTok Videos	58
5.4	Discussion of Results	62
6	Conclusions	67
6.1	Theoretical Contributions	67
6.2	Managerial Contribution	69
6.3	Methodological Contributions	70
6.4	Limitations and Future Research	72
	References	75
	Appendices	88
	Appendix 1 VVVA Analytical Protocol	88
	Appendix 2 Chartmetric Discovery Procedure	89
	Appendix 3 Anonymized Case Index & Metadata	90
	Appendix 4 VVVA Coding Manual (Matrices 1–4)	91
	Appendix 5 Data Management Plan	96
	Appendix 6 Explanation of the Use of AI	97
	Appendix 7 Row-Level Synthesis Tables (RQ1–RQ2 Audit Trail)	98

FIGURES

Figure 1. Textual hook and visual framing in a TikTok video (Lapid, 2021)	18
Figure 2. Russel's circumplex model of emotions (Adapted from Dzedzickis et al. 2020, 2)	28
Figure 3. Analytical Framework for Examining TikTok Post-Assemblages (S–St–R Heuristic)	34
Figure 4. Configuration framework of stimulus, staging, and response in short-form video content	66

TABLES

Table 1. Operationalization of Analytical Dimensions	41
Table 2. Analytical Dimensions of Micro-Storytelling Staging Configurations (Immediate Loops, Revealed Progressions, and Escalation Arcs)	55
Table 3. Cross-Layer Structure of Strategic Archetype Families (Stimulus, Staging, and Response Synthesis)	60
Table 4. Coding Manual Matrix 1 for Content Features and Visual Form	91
Table 5. Coding Manual Matrix 2a for Narrative Flow and Audio	92
Table 6. Coding Manual Matrix 2b for Progression Arc Configurations	92
Table 7. Coding Manual Matrix 3 for Content Messages, Characters, and Affect	93
Table 8. Coding Manual Matrix 4 for Engagement Response and Strategic Synthesis	94
Table 9. Configuration Synthesis and Engagement Modes by Post	99
Table 10. Content Messages and Staging Characteristics	102
Table 11. Payoff Timing and Resolution Cues	105
Table 12. Engagement Metrics and Dominant Modes for Matrix 4	107

1 Introduction

1.1 Relevance of the Study

Imagine uploading your track to a streaming service at midnight, knowing it joins 99,000 others released that same day—nearly 92 percent of them by independent artists just like you (Luminate 2024). Global distribution is more accessible than ever before, yet being heard is a different matter entirely: in 2023, 81 percent of artists on Spotify attracted fewer than 1,000 monthly listeners (Stassen 2024). Access, it turns out, is not the same as visibility. That distinction sits at the heart of this thesis and at the center of a digital music market so saturated that distribution and artistic talent alone no longer guarantee meaningful audience attention.

This condition, coupled with algorithmically governed visibility, produces what this thesis terms a *paradox of independence*: the very tools that free independent artists from traditional industry structures also subject them to new, less visible forms of control. Within the broader ‘show business’ ecosystem described by Danilova & Krupa (2021, 60), music operates simultaneously as a cultural product and an entertainment service, a duality that has long generated tensions between creative expression and commercial imperatives. For modern-day independent artists, these tensions have only intensified due to rapid technological change, shifting consumer behavior, and escalating competition (Haynes & Marshall 2018; Hodgson 2021; Frenneaux 2023; Radovanović 2024).

This paradox becomes most evident in the evolution of gatekeeping mechanisms within the music industry. Digital platforms promise artists the ability to bypass traditional industry intermediaries—namely, major labels—yet this same promise creates new dependencies on platform architectures for discoverability. The original gatekeepers have not vanished; they have transformed. Independent artists now compete in an attention economy shaped by platform-specific pressures (Haynes & Marshall 2018, 1974–1976; Frenneaux 2023, 126–128), making social media platforms key tools for audience engagement (Appel et al. 2020), music marketing, and artist brand communication (Danilova & Krupa 2021, 54–56).

Among these platforms, TikTok has emerged as a leading space for music promotion. Its recommendation algorithm, viral-trend culture, and organic reach have made it central to how independent artists seek visibility (Radovanović 2022, 52–63). TikTok content circulates through the ‘For You’ feed—an algorithmically curated stream where a post’s ability to engage determines its reach (Radovanović 2022, 60–61; Zhao & Wagner 2023; Jung 2024, 243–244). The continuous flow of content in the feed prioritizes platform-native posts that capture attention within seconds

(Mueller & Rajaram 2022, 5; Zhao & Wagner 2023), shifting the responsibility for visibility directly onto creators themselves (Frenneaux 2023, 126–127). Producing such content is not a secondary promotional task but a strategic form of communication on which an artist's discoverability depends (Frenneaux 2023, 125–129; Hodgson 2021, 2).

This dynamic is inseparable from the logic of engagement that underpins social media more broadly. Platforms are designed to foster ongoing relationships between creators and audiences (Gligor et al. 2019; Wang & Kim 2017), and content marketing has emerged as a key mechanism for sustaining those relationships (Tuten 2024, 227). In this thesis, engagement is operationalized through platform-visible behaviors, particularly higher-effort actions such as commenting, sharing, and saving relative to views (Muntinga et al. 2011, 15–16, 27–30; Malthouse et al. 2013, 272; Dolan et al. 2016, 265; 2019)—because within algorithmically curated feeds, these behaviors directly shape how the platform interprets and distributes content (Mueller & Rajaram 2022, 149; Tuten 2024).

Scholarship has examined how creators design posts to generate engagement (e.g., Lu & Shen 2023; Xiao et al. 2026). Yet much of this work still analyses the constituent elements of short-form video in isolation, focusing on audiovisual cues, persuasive strategies, or textual signals (Lu & Shen 2023; Zhang et al. 2025; Zhang & Zhang 2025). For independent artists navigating TikTok's rapid, interest-graph-driven feed without an established audience or major label infrastructure, this fragmentation is consequential. Each post must serve both as a creative expression and as an attempt to secure visibility. As a result, understanding how different components of post design contribute to engagement becomes an important part of explaining independent artists' promotional practices.

This master's thesis examines how independent artists design short-form video content on TikTok to drive engagement. It investigates how three analytically distinct dimensions of post design—how a post looks, what it communicates, and how it is structured—function within individual posts and shape visibility outcomes. The ambition is to illuminate the strategic patterns through which independent artists negotiate attention capture, brand expression, and the structural realities of platform distribution.

1.2 Research Gap

Several bodies of scholarship address the macro conditions shaping independent musicians' work in platform-mediated contexts, but they are rarely read in conversation with one another. Studies of the

music industry show how digital platforms have reshaped promotional practice, placing artists under what might be called an *engagement imperative*—a persistent pressure to cultivate algorithmic visibility at the expense of producing music (Haynes & Marshall 2018, 1975–1976; Edlom 2022; Frenneaux 2023, 125–126).

A separate line of work, rooted in artist branding and cultural consumption research, has long drawn attention to the affective and imaginative dimensions of audience relationships. Here, the argument is that fans do not simply consume music; they invest emotionally in the artists behind it (Hirschman & Holbrook 1982, 92–96). Building on this, scholars across media and communication studies have shown how parasocial interaction, perceived intimacy, and relational authenticity sustain artist–fan connections in networked spaces (Marwick 2013; Baym 2018; Stewart 2018; Sood & Rogers 2000; Danilova & Krupa 2021). These perspectives describe the structural pressures and relational expectations that artists navigate, yet they stop short of explaining how artists translate these conditions into communicative choices.

A second body of research approaches visibility from the perspective of content features that drive engagement. Empirical work in content marketing identifies structural qualities—vividness, interactivity, affective appeals—as antecedents of diffusion (Rowley 2008; Berger & Milkman 2012; de Vries et al. 2012, 85; Ashley & Tuten 2015; Vinerean 2017, 92–93; Cheung et al. 2020; He et al. 2021, 421–434). As TikTok has become a primary site for music discovery, this line of inquiry has been extended to platform-specific audiovisual attributes such as shot scale, text overlays, and video length (Bishop 2019; Chen et al. 2021, 7–8; Ling et al. 2022, 1–2; Zulli & Zulli 2022, 1879–1883; Cheng & Li 2024, 4–9; Jung 2024). This work identifies individual cues associated with visibility; how those cues operate in combination within the compressed temporality of short-form video remains a productive site for further inquiry.

Across these literatures, a productive analytical opening emerges. Each body of work contributes a distinct layer of understanding—the structural pressures artists face, the relational ideals that shape audience interpretation, the discrete features that predict engagement—yet the question of how creators organize these elements into coherent post-level arrangements remains largely unaddressed. The central process at stake is how creators assemble and time visual, textual, and sonic signals so that a short video is immediately interpretable to viewers unfamiliar with the creator, who often have no prior knowledge of the artist.

1.3 Purpose of the Study

The purpose of this study is to understand how independent artists design high-engagement TikTok short-form video content marketing strategies in an algorithmic environment and to interpret these strategies through an artist-branding lens. This leads to three guiding research questions:

1. What content features and communicative messages are present in independent artists' high-engagement TikTok short videos?
2. How do independent artists structure narrative meaning in these high-engagement videos through micro-storytelling sequences (hook–progression–payoff)?
3. What recurring strategic archetypes of artist-brand communication emerge across independent artists' high-engagement TikTok short videos?

Each short video is treated as a composite digital artifact that integrates audiovisual content, on-screen overlays, captions, hashtags, and measurable viewer engagement, enabling analysis of these elements as a unified communicative whole. Artist branding is analyzed through publicly available artifacts and the engagement they generate, maintaining an empirical focus on observable practice. The findings advance debates in artist branding and content marketing scholarship, particularly regarding how platform-specific conditions shape creative and strategic choices. Content features, messaging, and narrative structuring are thereby positioned as mutually constitutive elements of branding practice on algorithmically governed platforms.

1.4 Delimitations and Structure

This thesis focuses on independent artists working outside major-label contracts—a context in which visibility and brand communication largely depend on the artists' ability to produce short-form content tailored to the platforms they inhabit. TikTok was chosen as the sole platform because of its central role in the circulation of music-related short videos and its direct relevance to the concerns driving this research. Only posts from artists' official TikTok accounts are included, ensuring that the analyzed content directly reflects the artists' own communicative practice.

The primary unit of analysis is the *post-semblage*, which integrates the short video, its paratextual elements, and the resulting engagement traces. This operationalization specifies how the composite artifact is constructed analytically in the study. The content-centric scope is calibrated to the study's core analytical purpose: understanding how post-level design generates engagement. Accordingly, the analysis draws on publicly accessible posts and operationalizes engagement using publicly

visible interaction metrics. The empirical procedures, data selection criteria, and analytical instruments are detailed in Chapter 4.

The study moves from contextualization through to analysis and interpretation, with each chapter building on what precedes it. Chapter 1 establishes the research context, identifies the gap the study addresses, and sets out the research purpose and questions. This section concludes the framing by delimiting the scope. Chapter 2 then reviews the literature most relevant to the study, drawing on work in independent artist branding, platform visibility, and social media communication to lay the conceptual groundwork for what follows. Chapter 3 develops the theoretical framework, structuring the analysis around the key components of the *S-St-R* framework.

Chapter 4 addresses methodology: it outlines the philosophical orientation underpinning the study, describes the research design, specifies the data sources and unit of analysis, and introduces the analysis instrument developed for this purpose. Questions of trustworthiness and ethics are also discussed there. Chapter 5 presents the empirical findings and interprets them in light of the research questions and the studies reviewed in Chapters 2 and 3. Chapter 6 concludes the thesis by consolidating the study's contributions, drawing out practical implications, and identifying directions for future research.

2 Independent Artist Branding in the Algorithmic Era

2.1 Independent Artists under Algorithmic Gatekeeping

Independent artists now make up a significant and growing share of the recorded music market. For the purposes of this thesis, ‘independence’ refers to musicians operating outside the major-label system who retain control over their creative output, distribution, and branding. In 2024, the independent sector accounted for roughly 29.7 percent of global recorded music distribution (Tadesse 2025). While this independence offers artists greater creative autonomy, it also places them in an intensely competitive digital marketplace characterized by a vast, continuously expanding supply of content (Frenneaux 2023, 125–131; Murphy & Hume 2025, 175–178). Digital distribution has produced a persistent oversupply of music, where the majority of artists attract only very small audiences (Luminate 2024; Stassen 2024). The result is audience saturation, positioning visibility as a central and ongoing challenge for independent creators (Danilova & Krupa 2021, 51).

The difficulty of standing out is closely linked to modern gatekeeping practices in the music industry. In the past, major record labels controlled which artists could reach audiences and enter the market. Today, that gatekeeping role has shifted to platforms such as streaming services and social media networks. (Bishop 2005, 443–445; Lewis et al. 2005, 349; Murphy & Hume 2025, 174.) Sites such as TikTok and Spotify now shape discoverability through a combination of algorithmic ranking and editorial curation. Bonini and Gandini (2019, 1–7) describe this as “*algorithmic*” gatekeeping. These systems have been criticized for reinforcing existing engagement hierarchies, privileging artists with greater visibility or promotional resources, and thereby constraining the capacity of lesser-known creators to attain meaningful reach (Haynes & Marshall 2018, 1985–1989; Frenneaux 2023, 125).

Algorithmically governed visibility has reshaped what it means to be an independent musician on a day-to-day basis (Haynes & Marshall 2018, 1974, 1981, 1990; Hodgson 2021, 2; Frenneaux 2023, 125). As discoverability increasingly depends on platform-driven content circulation, the boundary between creative production and promotional activity becomes more porous (Bonini & Gandini 2019, 1; Hodgson 2021, 2; Frenneaux 2023, 125). While navigating market competition has always required artists to develop new business competencies, the current digital context intensifies this convergence, requiring independent musicians to operate at once as creators, marketers, and entrepreneurs within a highly competitive, visibility-oriented ecosystem. (Frenneaux 2023, 125–129.)

Danilova and Krupa (2021, 60) clearly describe two areas of the music business: ‘pure music,’ which concerns inspiration and composition, and the commercial side, which focuses on making money. On social media, the commercial side pressures artists to adjust their promotional strategies and, often, their creative choices in direct response to platform logic. Many independent musicians report finding marketing genuinely difficult (Radovanović 2024), which points to a broader structural condition: success in the music market depends on creative excellence as much as on the capacity to produce content that aligns with the operational logics of platform-specific visibility (Hodgson 2021, 2).

Cutting across both algorithmic and human curation (Bonini & Gandini 2019, 1) is the expectation of authenticity: the drive for visibility is inseparable from the need to appear genuine. Independent musicians must navigate the competing demands of sincere self-expression and strategic self-promotion, often prioritizing building audience trust over overt commercial messaging. (Stewart 2018, 70; Danilova & Krupa 2021, 62.) On social media, authenticity is less a fixed personal quality than something actively constructed—performed through communicative cues that signal familiarity, openness, and relatability (Sood & Rogers 2000, 388).

To be an independent artist is to operate within a set of competing imperatives—creative autonomy, algorithmic visibility, and sustained audience connection—each of which exerts a distinct and ongoing pressure on artistic practice. This thesis conceptualizes this as the *paradox of independence* as introduced in Chapter 1. Rather than choosing definitively between authenticity and commercial self-promotion, independent artists must continually renegotiate this boundary, maintaining credibility while adapting to the visibility logic imposed by digital platforms.

2.2 Artist Branding on Social Platforms

In social media contexts, the *paradox of independence* is often negotiated through artists’ branding practices. In the independent sector, artist branding is shaped by the self-managed cultivation of digital presence and the ongoing relational exchanges between artists and audiences online (Hracs 2012, 456; Baym 2018, 163–164). On platforms such as Instagram and TikTok, this process becomes participatory in practice: audiences engage in continuous interaction by sharing, commenting on, and circulating content, thereby contributing to how an artist’s identity is constructed and perceived (see Baym 2018, 163–164; Haynes & Marshall 2018, 1974–1976; Nickell 2020, 52; Radovanović 2024). In this context, authenticity functions at once as an audience expectation and a strategic performance through which artists signal credibility and relatability (Marwick 2013, 7, 119, 207; Nickell 2020, 52–59; Arriagada & Bishop 2021, 568–569).

To conceptualize how artists construct their brands on digital platforms, this study draws on Danilova and Krupa's (2021) perspective on brand management alongside Stewart's (2018) model of audience engagement. In branding theory, a brand constitutes a constellation of recognizable signs and symbols through which audiences identify and differentiate an entity, whether a company or a musician (Danilova & Krupa 2021, 54, 63). On social media, such symbolic identities take shape through recurring stylistic patterns—visual framing, aesthetic choices, and thematic consistency across posts—that together form a coherent visual brand identity (Phillips et al. 2014, 226). Repeated exposure to these cues fosters audience recognition and deepens the emotional and relational bonds between audiences and the artist (Fournier 1998, 365; Chen et al. 2015, 100). Supporting this perspective, Stewart (2018, 52–58) identifies two primary drivers of engagement in music contexts relevant here: audience responses to the artist as a person and responses to the music itself.

Audiences frequently develop perceived relationships of familiarity or intimacy with mediated personalities, even when such relationships remain one-sided (Sood & Rogers 2000, 387–388). On digital platforms, these parasocial relationships are reinforced through communicative cues that signal accessibility—personal storytelling, direct address to the viewer, or eye-level camera framing (Stewart 2018, 70). This thesis treats parasocial interaction as an analytical lens for examining how intimacy and relational positioning are constructed in short-form video content and how these cues shape audience interpretations of an artist's persona.

Music consumption also tends to involve hedonic and experiential motivations rather than purely utilitarian ones (Hirschman & Holbrook 1982, 94). Listening to music can provoke emotional responses, memories, and imaginative associations, which in turn influence how audiences engage with artists and their creative output (Lacher 1989, 371; Lacher & Mizerski 1994, 368; Stewart 2018, 54–55). Consumers seek emotional gratification and identity affirmation through cultural consumption, responding most strongly to messages that relate to their self-concept or reinforce their self-esteem (Vaughn 1980, 30; Batra & Ahtola 1990, 159; Johar & Sirgy 1991, 24, 26). In short-form video contexts, expressive elements such as vocal delivery, lyrical excerpts, and visual performance cues are therefore key to how audiences experience and interpret the artist's brand. Prior research thus indicates that artist branding is better understood as an evolving pattern of communicative signals rather than a fixed identity. Examining audiovisual elements allows tracing how branding constructs such as persona, intimacy, and value alignment are articulated within individual posts.

2.3 TikTok's Algorithmic Visibility

Visibility on TikTok cannot be reduced to creator intent or the formal qualities of their content. Rather, it takes shape through the interplay among platform infrastructures, audience engagement, and users' creative practices, reflecting the sociomaterial inseparability of the social and the technical in digital media systems (Orlikowski & Scott 2008, 434, 455). From this perspective, agency is distributed across people and technological systems alike—an artist's visibility on TikTok is, in effect, co-produced through the convergence of creative decisions, audience engagement, and algorithmic mediation.

Within this sociotechnical setting, users have developed a distinctive vernacular of evolving conventions, genres, and communicative styles (Gibbs et al. 2015, 257–258). On TikTok, this vernacular is evident in practices such as short-form storytelling, trend participation, and remixing sounds that align with the platform's native modes of interaction (Jung 2024, 243–256). For independent artists and marketers, familiarity with this vernacular is essential. Promotional content gains traction when it is embedded in formats that audiences read as genuine participation—content that conforms to the platform's native communicative conventions. TikTok, therefore, functions less as a conventional social networking site and more as a cultural marketing space in which technological design, creative expression, and strategic self-presentation are difficult to disentangle (Mueller & Rajaram 2022, 21–22, 24–25; Tuten 2024, 11, 21, 149, 240; Jung 2024, 247, 253–254).

Music is a foundational element of TikTok's infrastructure. As a mobile-first, short-form multimedia platform built primarily around entertainment, TikTok enables users to produce and share audiovisual content through platform-integrated tools and trending audio formats (see Jung 2024, 244; Tuten 2024, 265; Ceci 2025). Users can layer music, text, visual effects, and editing functions within a single interface, producing content quickly and in forms that fit the platform culture (Jung 2024, 244–245). This multimodal layering is illustrated in Figure 1, which demonstrates how these components are combined within a single TikTok video frame.



Figure 1. Textual hook and visual framing in a TikTok video (Lapid, 2021)

The figure illustrates how musical performance, on-screen text, and visual presentation are combined within a single TikTok video frame. The opening text functions as a hook that invites audience interpretation, while the co-presence of lyrics and close-up framing demonstrates how multiple semiotic elements are layered to capture attention and construct meaning in platform-native formats. The example features Lyn Lapid, who initially gained visibility on TikTok as an independent creator before later signing to a major label (Republic Records).

TikTok's user base expanded during the social disruptions of the COVID-19 pandemic, significantly accelerating an already rapid expansion, and it soon ranked among the most downloaded applications worldwide (Radovanović 2022, 52; Vizcaino-Verdu & Abidin 2022, 884). As its audience grew, the platform's influence on how audiovisual media is created, consumed, and circulated became increasingly pronounced—shaping how creators produce content, how audiences engage with songs, and how the music industry identifies and promotes artists. (Radovanović 2022, 52.) Research further supports TikTok's influence in the music industry. Caswell and Daniels (2022) found that a quarter of artists who appeared on the US Spotify charts between 2020 and 2021 had initially built their visibility on this platform, with many (46%) later signing to major record labels.

Central to this process is TikTok's recommendation system. The platform translates user behavior into data that determines which videos surface in the users' 'For You' feed (Zulli & Zulli 2022, 1878). Technically, the system draws on both collaborative and content-driven filtering to predict what individual users are likely to engage with (Martins et al. 2020, 1–2; Zhou 2024, 205). The former identifies patterns across viewing histories and engagement behavior; the latter evaluates features embedded in the videos themselves, such as hashtags, captions, audio tracks, and visual attributes (Zhou 2024, 205). The result is feeds that are highly personalized and continuously updated in response to user feedback (Portugal et al. 2018, 205–206; Zhou 2024, 205; Salles 2025, 6–8). From a sociomaterial perspective, the algorithm can be understood as a kind of non-human audience—one that reads creative content through data signals and determines which posts circulate.

Although the algorithm's internal logic remains largely opaque, its effects become legible through engagement metrics and patterns of viral circulation. Viral content spreads through networks of interconnected users via contagion-like processes, in which ideas and messages travel from one individual to another through social ties. For marketers, virality is appealing precisely because it enables the widespread dissemination of messages at relatively low cost, potentially reducing reliance on resource-intensive advertising campaigns. (Curmi 2022, 146–151, 157–158.) Two characteristics are considered especially defining for viral content: the speed of transmission and the sudden engagement spikes that dramatically extend a message's reach (Han et al. 2020, 579). A widely cited example of viral diffusion is the success of Lil Nas X's song 'Old Town Road.' The track gained popularity on TikTok through the 'Yee-Haw Challenge,' where users created short videos featuring cowboy-themed transformations. As the trend spread across the platform, the song gained mainstream visibility, eventually leading to a remix with Billy Ray Cyrus and a record-breaking 19-week run at number one on the Billboard Hot 100. (Alexander 2019; Chow 2019.)

TikTok's architecture promotes the circulation of viral content through collaborative features such as Sound, Duet, and Stitch, which enable users to engage with existing videos and help build chains of audiovisual creation (Radovanović 2022, 55–61; Savic 2021, 3184). Among these, sound occupies a particularly central role. Users may either create original audio or reuse sounds from other videos, enabling short audio fragments—whether music, speech, or sound effects—to travel across thousands of posts (Radovanović 2022, 55). Sound thus operates at once as a cultural medium and a technical connective link, binding dispersed videos into shared trends and challenges.

The Duet feature further facilitates collaborative participation by allowing users to create split-screen videos alongside an existing post (Radovanović 2022, 56; TikTok n.d.-b). Through successive responses, a single video can generate extended chains of collaborative reinterpretation. One example is Nathan Evans's viral performance of the sea shanty 'Wellerman,' which inspired users to contribute additional vocals and instrumental layers through duet performances, creating a collective musical production (Radovanović 2022, 57). Like Duet, the Stitch feature enables users to incorporate short segments of other videos into new posts, supporting commentary, reinterpretation, and reaction-based content (Radovanović 2022, 60; TikTok n.d.-c).

TikTok's emphasis on brevity shapes audiovisual narratives designed for immediate visibility. Historically featuring content ranging from 15 seconds to 3 minutes, the platform creates an 'accelerated, hypnotic pace' that demands creators capture viewers' attention within the first few seconds. (Radovanović 2022, 60–61; Jung 2024, 243–244.) In music production, an informal convention has emerged in which artists tailor songs by foregrounding verses or choruses that fit short video lengths, placing hooks closer to the opening to serve as platform discovery tools (Radovanović 2022, 61). These shifts point to how cultural production is increasingly governed by algorithmic logic and the need for content to be recognizable through coded metadata (Morris 2020, 2–4).

Consequently, for independent artists, branding unfolds where platform infrastructures and audience participation meet. On TikTok, algorithmic recommendations, remix culture, and short-form conventions together shape how artists become visible. Creators must stay recognizable and authentic to audiences while producing content that remains legible to algorithms that privilege novelty and engagement. The framework for examining how they do so is developed in next Chapter 3.

3 Design of Short-Form Content in Algorithmic Platforms

3.1 Engagement as Observable Response Traces

On TikTok, the relational orientation between artists and audiences described in Section 2.2 becomes visible only through a narrow set of platform-mediated behaviors. In social media marketing¹, audience responses are typically measured through highly visible indicators such as views, likes, and follower counts (Rogers 2021, 271). Scholars have long criticized these measures for producing a distorted, often performative picture of audience interaction, even if they retain baseline branding value (Lipschultz 2020, 10, 12; Rogers 2021, 272–275). Empirical studies reinforce this critique; for instance, Tafesse and Wood (2021, 6) demonstrate that large follower counts do not guarantee marketing effectiveness and are often negatively associated with engagement rates.

In response, research has increasingly shifted toward behavioral traces that signal active user investment rather than passive exposure (Rogers 2021, 276–277; Tafesse & Wood 2021, 3). This thesis adopts the concept of Social Media Engagement Behaviors (SMEB), defined as user actions on social platforms driven by specific psychological motivations (Dolan et al. 2016, 265). SMEB spans a continuum from passive consumption to effortful participation, with prior research consistently identifying higher-effort behaviors as more reliable indicators of relational or identity-relevant engagement (Muntinga et al. 2011, 15–16; Malthouse et al. 2013, 272). This distinction is particularly salient in short-form video contexts, where most users remain passive viewers while only a minority engage in effortful, platform-visible actions (Van Dijck 2009, 44). What an independent artist’s post must do, then, is provoke precisely that minority response—and understanding what provokes it begins with the content itself.

This emphasis on observable, effortful behaviors is consistent with relationship marketing theory, which conceptualizes engagement as value-generating interaction within ongoing relationships between brands and audiences (Morgan & Hunt 1999, 281; Shay & van Der Horst 2019, 27). From this view, social media users are not passive recipients but active participants in processes of co-creation, identity expression, and meaning-making (Jensen Schau & Gilly 2003, 385, 397; Cheung & To 2016, 265; Akdevelioglu & Venkatesh 2022, 60–63). Accordingly, engagement reflects forms

¹ F. Li et al. (2021, 54) define social media marketing strategy as an integrated set of activities that transforms social connections and interactions into strategic marketing outcomes, grounded in users’ motivations and deliberate engagement initiatives.

of involvement that contribute to relational development and emotional attachment (see Sinha et al. 2011, 189, 194–195; Sashi 2012, 264; Chahal et al. 2020, 192–193; Hanlon & Tuten 2022, 10).

Within this framework, the study operationalizes active engagement through three effortful actions, each associated with a distinct motivational orientation. Commenting typically signals a desire for interaction, visibility, or self-expression (Muntinga et al. 2011, 29–30). Sharing indicates that the user perceives the content as valuable enough to circulate within their own networks, functioning as a form of endorsement (Dolan et al. 2016, 267). Saving reflects perceived ongoing relevance, as users curate content for future use or inspiration (Van Dijck 2009, 44; Muntinga et al. 2011, 27–28). Together, these three behaviors constitute the response layer of the analytical framework—and the question of what stimulus produces them is where the analysis turns next.

3.2 Content Stimulus Building Blocks

If the three behaviors identified in Section 3.1 constitute the response layer of the framework, the analytical question that follows is what, at the level of the post itself, produces them. Empirical studies confirm that multimodal content cues—visual presentation, audio design, and textual overlays—shape engagement behaviors, including commenting and sharing (Tao et al. 2026; Xiao et al. 2026, 154–156). In such contexts, posts function as artifacts that present these cues to viewers at the moment of encounter, embedded within a continuous stream of competing content.

Examining the stimulus level means asking what is actually in the post—and how its constituent elements are organized to make a response not just possible but likely.

To enable systematic analysis, prior research distinguishes between two analytically separable but co-present dimensions: observable content features and interpreted content messages. Adapting from Han et al. (2020, 578, 580–582), *content features* refer to perceptible formal elements that viewers can directly observe, including visual composition, proxemics, text overlays, editing patterns, and audio sourcing. *Content messages*, by contrast, are the meanings that emerge from the interaction of these elements within a multimodal whole (see also Zhang & Zhang 2025; Zhang et al. 2025, 2–3). Although analytically separable, these dimensions are empirically intertwined; treating them as distinct layers makes it possible to trace, across cases, how particular formal arrangements come to carry particular meanings.

Situating these content characteristics within the broader development of marketing communication clarifies why content has become a central analytical unit. Tuten (2024, 225) defines content as a social object that encourages participation and constitutes the primary unit of value on social media,

implying that value must be enacted within individual posts and extend beyond abstract brand identity. More broadly, content marketing refers to the creation and distribution of valuable and relevant material aligned with a brand's identity to attract audiences and stimulate action (Content Marketing Institute n.d.; see also He et al. 2021, 420–421; Tuten 2024, 227). On continuously scrolling platforms, communicative value depends on how effectively the post is made interpretable through embedded cues.

Brand communication in social media contexts is frequently conceptualized through constructs such as brand personality, personification, and brand personas, which describe how organizations cultivate distinctive identities and position themselves as relational actors with respect to audiences (Tuten 2024, 194). Organizations curate their presence to project recognizable personas that differentiate them from competitors and enable the brand to function as a relational partner in ongoing consumer–brand relationships (Fournier 1998, 344). Credible persona presentation has been linked to higher audience engagement (Chen et al. 2015, 97–98). Rather than treating persona as a fixed identity that precedes communication, it can be understood as emerging through recurring communicative cues that construct recognizability, relational stance, and character within individual posts.

This competition for attention unfolds within platform conditions that shape how meaning is interpreted across contexts. TikTok's recommendation system distributes content primarily through an interest graph rather than a follower network, allowing posts to reach audiences beyond the creator's established community (Eadon 2024, 3). The result is what Palombo (2024, 95–105) describes as *context collapse*—a condition in which distinct audience groups with divergent norms and expectations are merged into a single, undifferentiated audience, thereby exposing content to viewers for whom it may be inappropriate or be interpreted differently. To make this concrete: An independent artist posting a self-deprecating joke about music pressures may find the same post, shared by the algorithm, shown to tens of thousands of unfamiliar viewers who lack context. To them, irony might seem sincere, humor as complaint, or the persona as unfamiliar. The same content, seen by different audiences, can have conflicting interpretations and readings. Under these conditions, algorithmic distribution on platforms such as TikTok may result in content being encountered by unintended audiences, increasing the likelihood of misinterpretation and conflicting evaluative responses.

From a stimulus-level perspective, interpretability is an outcome of how multimodal cues are organized within a post—what is made perceptible at first encounter, before context or creator

intent can be inferred. The five feature categories and three message dimensions introduced in the subsections that follow are the analytical instruments through which that organization becomes legible.

3.2.1 Content Features

Content features are a set of objective, easily identifiable cues and structural signals—observable properties of the post that guide how attention is allocated and how content is interpreted. Five such feature categories are relevant here, moving from the most immediate first-contact properties a viewer registers to those that require more inferential work: *temporal structure*, *visual composition*, *platform-specific aesthetic conventions*, *paratextual packaging*, and *audio infrastructure*. Each operates differently. Each matters. Video length is perhaps the most deceptively simple of these. TikTok technically allows videos ranging from 1 second to 60 minutes (Breitman 2024; TikTok n.d.-a), yet the platform remains optimized for short-form media (Hutchinson 2022; Jung 2024, 243–244). Empirical work supports the point: shorter TikTok videos tend to generate higher engagement (Chen et al., 2021). Concise formats appear to align with the fragmented viewing habits the platform both reflects and reinforces.

Visual grammar is the second feature category, concerned with how viewers are positioned in relation to what they see. Drawing on Kress and van Leeuwen's (2021) theory of visual communication offers a useful framework here: camera angles, framing, and spatial relations are not neutral choices but contribute to the construction of social meaning. A front-facing angle invites direct viewer involvement; a side angle introduces distance. Looking up at a subject may connote authority, while eye-level framing tends to communicate equality. On vertically oriented mobile screens, visual composition often concentrates the subject within a narrow frame, foregrounding facial expressions and bodily gestures (Mustikawati et al. 2023, 129–130)—a constraint that, on TikTok at least, appears to intensify rather than limit communicative possibility.

Creators on the platform frequently combine medium and close-up shots with eye-level framing, a pairing that reduces perceived social distance and fosters a sense of proximity between maker and viewer (Zummo 2022, 101–102). This resonates with the older literature on parasocial interaction, which describes how audiences come to experience mediated performers as familiar and intimate figures despite the one-sided nature of the relationship (Horton & Wohl 1956, 215–216; Sood & Rogers 2000, 389). Direct camera gaze intensifies this further: Kress and van Leeuwen (2021, 104–155) call these 'demand' images, where the depicted participant looks directly at the viewer and, in a sense, addresses them. That such choices carry measurable consequences is supported by research

indicating that close-up framing and second-person camera perspectives are associated with higher levels of circulation and engagement on TikTok (Ling et al. 2022, 7–8; Cheng & Li 2024, 210–211).

Visual composition, then, shapes relational positioning before a viewer has processed what a video is about. The platform’s distinctive vernacular aesthetic—what Jung (2024) calls the widely recognized ‘TikTok look’—operates at a different register, one concerned less with spatial orientation than with tone, credibility, and cultural belonging. This encompasses playful, often amateurish visual styles: app filters, intimate everyday settings, and dense multimodal combinations of text, sound, and image. It would be a mistake, however, to read these stylistic cues as signs of limited production capacity or carelessness. Scholarship on social media creators has consistently associated such aesthetics with deliberate performances of spontaneity, authenticity, and relatability (Abidin 2016, 6, 13; 2017, 7–8; Serazio 2024, 29). Abidin (2017, 7–8) captures this with the term ‘calibrated amateurism’—the deliberate staging of rawness to sustain a sense of intimacy and credibility with followers. Creators therefore move strategically between polished and informal registers as part of ongoing impression management (Abidin 2016, 10; 2017, 7–8), and the vernacular look becomes less an aesthetic default than a communicative resource through which audiences read tone, credibility, and creative identity.

Where vernacular aesthetics shape how a post is read tonally, paratextual and metadata elements constitute a fourth, often underappreciated layer of observable features that shape how it is read categorically. Captions, hashtags, tags, and text overlays do more than describe content: they orient viewers by signaling topics, moods, or communities and, at the same time, position posts within the platform’s systems of categorization and circulation. Bishop’s (2019, 2589–2590) concept of ‘algorithmic gossip’ is useful here, as it describes how creators collectively develop informal theories about algorithmic visibility and share strategies to improve reach. The concept was formulated in relation to YouTube. Nevertheless, the broader pattern Bishop identifies—creators experimenting with and exchanging knowledge about hashtags, timing, and engagement tactics to improve visibility—appears to hold across platforms (Cotter 2019, 902–903). Metadata, in this sense, functions not just as description but as a strategic resource through which posts are aligned with platform-specific distribution logic.

The fifth feature—audio—was already discussed at the platform architecture level in Section 2.3, where its role in collaborative circulation was outlined. At the stimulus level, it is worth examining more closely, as audio is arguably TikTok’s most distinctively participatory content feature. Sound

operates on at least two registers at once: as a technical layer embedded in the post, and as a medium for remix and intertextual connection across videos. Shared audio allows users to reuse and reinterpret existing sounds, linking otherwise separate posts through shared sonic cues (Pilipets 2021, 111–112; Bhandari et al. 2022, 5). Empirical support for audio’s centrality comes from Zhu et al. (2020, 7), whose content analysis found that most TikTok videos contained background music—often original music created by the uploader—and that subtitles were also frequently used. The study’s focus on health communication accounts is a limitation worth noting, but the findings nonetheless illustrate how audiovisual and textual cues are layered in TikTok’s short-form videos.

These five bundles of content features—temporal structure, visual composition, aesthetic vernacular, paratextual packaging, and audio infrastructure—constitute the stimulus-level resources through which TikTok posts become legible within a saturated, algorithmically curated feed. None of them operates in isolation; it is their combination and the tensions between them that shape how quickly a post registers, how viewers orient toward it, and how it circulates. That combination is what the next section makes analytically tractable. Knowing what is in a post, however, is not yet the same as knowing what it means—and that is where the analysis turns next.

3.2.2 Content Messages

On a platform where posts reach viewers with no prior knowledge of the creator, meaning must be constructed rapidly and from limited cues—emerging through the interplay of visual imagery, sound, written text, and bodily performance that together form an integrated communicative whole (Kress & van Leeuwen 2021, 17; Fazeli et al. 2023, 2–6; Grzenkowicz & Wildfeuer 2025, 1143–1145). The message level is where that construction becomes analytically legible (Han et al. 2020, 577).

In such settings, viewers rely on recurring cues to stabilize meaning during content consumption. Drawing on prior research on online virality, TikTok video features, and multimodal communication, this study conceptualizes three recurring dimensions through which audiences commonly orient their interpretation of short-form videos: *topic*, *vernacular genre*, and *affect*. (see Hansen et al. 2011, 34; Han et al. 2020, 577; Ling et al. 2022, 6–8; Grzenkowicz & Wildfeuer 2025, 1144.) These dimensions function as interpretive anchors—allowing viewers to infer what a video is about, how it should be read, and what emotional register it occupies, even when the content originates outside their familiar social or cultural context (Han et al. 2020, 577; Grzenkowicz & Wildfeuer 2025, 1146–1147).

The topic refers to what the post is about and, therefore, provides the conceptual entry point for interpretation. In social media research, content messages often include elements such as humor, affective cues, opinions, and themes, including politics and popular culture, that audiences interpret when engaging with posts (Taecharungroj & Nueangjamnong, 2015, 288–289; Han et al., 2020, 577, 581). This is particularly relevant on algorithmically curated platforms, where recommender systems determine which videos become visible to different audiences, thereby shaping how content circulates across the platform (Bishop 2019, 259; Chen et al. 2021, 2–3).

Another way viewers orient themselves is through vernacular genres—recurring platform-specific formats that emerge from the interaction between platform affordances and user practices. Social media platforms develop distinctive combinations of styles, grammars, and conventions that shape how communication is produced and interpreted. (Gibbs et al. 2015, 257.) On TikTok, these conventions manifest in recurring video formats such as lip-sync performances, dance routines, comedic sketches, life-sharing videos, and makeup tutorials (Ling et al. 2022, 6). These formats structure participation by providing recognizable templates that can be imitated, remixed, and adapted, enabling content to circulate through processes of imitation and replication (Abidin & Kaye 2021, 59–60; Zulli & Zulli 2022, 1873).

Where topic and vernacular genre orient the viewer conceptually and structurally, affective orientation operates at a more immediate register—the emotional tone a video conveys, whether through excitement, humor, calm, nostalgia, or melancholy. Emotional meaning in media communication is analytically complex, encompassing overlapping constructs such as feelings, moods, and evaluative preferences (Scherer 2005, 695). A widely used framework for mapping this terrain is the circumplex model of affect (Russell et al. 1980, 1164), shown in Figure 2, which organizes emotional states along two dimensions: valence (positive–negative) and arousal (high–low). Later research has applied this model to digital media contexts to examine how emotional tone shapes audience interpretation and interaction (Dzedzickis et al. 2020; Schreiner et al. 2021). In this thesis, the circumplex model provides the conceptual basis for interpreting the affective orientation communicated in TikTok videos.

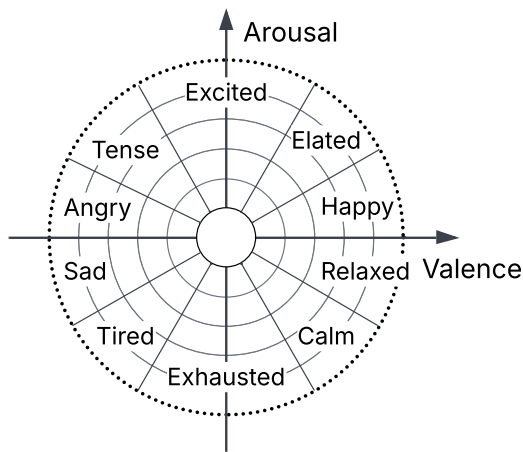


Figure 2. Russel's circumplex model of emotions (Adapted from Dzedzickis et al. 2020, 2)

Emotional value also shapes how digital content circulates (Pivecka et al. 2022, 1–2). Studies in marketing and communication show that content eliciting high-arousal emotions is more likely to be shared than content associated with low-arousal states (Berger & Milkman 2012; Nelson-Field et al. 2013, 210–211; Li et al. 2015, 6–7; Pivecka et al. 2022, 5; Rodas-Coloma et al. 2026, 12–13). High-arousal emotions activate behavioral responses and increase the likelihood that audiences pass content along to others (Li et al. 2015, 6–7). Research on viral video advertising likewise indicates that emotional responses—particularly emotional arousal—can influence viewers' motivations to share advertising content on social media (Nikolinakou & King 2018, 721–722).

Content messages illuminate what a post communicates at the moment of encounter—its subject, its generic register, its emotional pitch. What they do not yet account for is how that meaning unfolds across the seconds a viewer actually spends watching. On platforms defined by high velocity, meaning must be organized temporally, and audiences need to recognize, follow, and resolve it almost instantaneously.

3.3 Narrative Staging in Short-Form Video

Meaning in TikTok posts is shaped by how multimodal signals are organized in time: the ways creators sequence actions, cues, and events so that viewers can read them as a coherent narrative moment. The temporal pressure on this process is considerable. Research shows that users decide within the first seconds whether to keep watching (Ma & Jiang 2024, 1–2; Yang et al. 2024, 117; David & Roberts 2025, 1293), which pushes creators toward unusually compact narrative forms.

Mueller and Rajaram (2022, 149) describe one such form as *micro-storytelling*—brief, self-contained sequences built around a single situation, action, or moment.

Drawing on prior research, this thesis synthesizes recurring narrative patterns of these *micro-stories* into three analytically distinct stages: *a hook, a progression, and a payoff*. This structuring reflects the conditions of attention on short-form platforms, where viewer engagement is highly contingent and content must establish relevance within the opening seconds to sustain participation (Li et al. 2025, 294–295; Salles 2025, 9–10; Xiao et al. 2025, 678–679). Creators adapt by organising content into temporally compressed sequences that establish relevance immediately and sustain interpretability throughout the post. Accordingly, the hook functions to secure initial attention by presenting recognizable cues and establishing the audiovisual register of the content at the moment of encounter (Nastisin et al. 2025, 64–65; Yang et al. 2025, 1, 7). The progression then develops this initial orientation by unfolding cues, actions, or information across modalities, maintaining viewer involvement while incrementally structuring meaning (Fazeli et al. 2023, 8; Grzenkiewicz & Wildfeuer 2025, 1144). Finally, the payoff provides a moment of closure—such as a reveal, a punchline, a rhythmic accent, or a visual transformation—that retrospectively organizes the sequence into a coherent and purposeful whole (Bösch & Divon 2024, 5089). Together, hook, progression, and payoff describe how storytelling operates as a mechanism for capturing, sustaining, and resolving viewer attention within highly competitive feed environments.

What makes condensed narrative legible is the presence of identifiable agents and some evaluative point—a reason, however minimal, for the sequence to matter (Adaval & Wyer 1998, 210; Schank 1999, 89–97; Polletta et al. 2011, 111; Mills & John 2025, 36). Classical storytelling models describe a movement from an initial situation through a complication to a transformed state (Labov 1972, 362–370; Mills & John 2025, 36–38), but short-form content does not always follow this arc in full. Drawing on Labov’s (1972, 362–370) model of narrative structure, Bouvier & Rasmussen (2022, 109–111, 129) demonstrate how stages such as orientation and complicating action may be rearranged or replaced by brief previews in short-form clips (see also Ledin & Machin 2018, 147). TikTok staging, then, frequently works through selective foregrounding: only fragments of a storyline appear, yet viewers reconstruct a coherent narrative effect from them.

The brevity of TikTok videos creates a specific analytical problem: characterization, emotional tone, and relational positioning—elements that longer formats can develop gradually—must be established almost immediately. That is not a limitation. It is the condition under which short-form meaning is made. In this study, these are treated as staging practices that invite viewers to read the

narrative situation. This framing follows prior work emphasizing that narratives depend on identifiable agents whose actions make events intelligible (Escalas & Stern 2003, 566–567; Polletta et al. 2011, 112–113; Mills & John 2025, 38–39). Under strong temporal compression, characterization tends to surface through minimal expressive cues—a facial expression, a gesture, a shift in tone, or a brief situational frame—from which viewers infer social characteristics, intentions, and relational orientations almost instantaneously. (Ambady & Rosenthal 1992, 256–257; Murphy & Hall 2021, 1–2.)

Affective cues operate in a comparable but distinct register. Where affect at the message level establishes the emotional tone of a post as a whole, affective cues within the staging sequence function temporally, intensifying or redirecting how the unfolding situation is perceived. Humor, surprise, and nostalgia are among the signals that can make individual moments stand out and amplify narrative development in condensed video (Schreiner et al. 2021, 335; Ameh & Emmanuel 2025, 1–2).

Direct address, expressive camerawork, and intimate framing can produce a comparable effect at the relational level, generating a sense of closeness between performer and viewer. Two concepts from narrative research are relevant here: narrative transportation—cognitive and emotional immersion in a story world (Green & Brock 2000, 701–703)—and parasocial interaction, which describes the perceived intimacy audiences develop with mediated figures (Zhu et al. 2024, 108–110; Mohammed et al. 2025, 209). Both concepts explain why the *hook–progression–payoff* sequence works: a viewer drawn into a story world and positioned as a familiar intimate is far more likely to follow the sequence through to its payoff.

Short-form staging is also closely connected to networked participation. The temporal organization of micro-narratives often incorporates what Divon and Ebbrecht-Hartmann (2025, 3–8) describe as playability, a form of semiotic openness that allows viewers to inhabit, remix, or extend a scenario. The POV (point-of-view) format is a clear example: it presents a structured perspective that is nonetheless open enough to invite imaginative projection from a wide range of viewers. This participatory logic extends further through the platform affordances discussed in Section 2.3—Duets and Stitches in particular—which distribute narrative construction across multiple creators and embed individual posts within broader conversational chains. A post that is well-sequenced enough to be legible and open enough to be remixable is one the platform’s distribution logic is inclined to carry forward.

Short-form staging, in other words, is not a reduced version of longer narrative forms but a distinct practice shaped by temporal pressure, platform architecture, and participatory convention. The mechanisms examined here—*hook–progression–payoff* structures, quick character development through minimal expressive cues, and open participatory templates—are strategic tools by which creators manage attention, meaning, and emotion on their own terms, within a form that has developed its own logic and standards of coherence.

3.4 Configuration Logic for Integrating Stimulus, Staging, and Response

An independent artist uploading a post to TikTok is making three kinds of decisions simultaneously: what to include in the video, how to organize it over time, and what response to elicit from the viewer. The preceding sections have developed the theoretical vocabulary for each of these layers in turn. This section brings them into relation. The purpose of this study is to understand how independent artists design high-engagement TikTok short-form video content marketing strategies in an algorithmic environment and to interpret these strategies through an artist-branding lens. Drawing on an adaptation of the stimulus–organism–response (S–O–R) paradigm (Mehrabian & Russell 1974), the framework is conceptualized here as a *stimulus–staging–response* (*S–St–R*) logic—an analytical structure for examining how the observable elements of a short-form post cohere under sociotechnical conditions to generate engagement.

Two foundational assumptions underpin the analytical approach. The first concerns platform distribution, which is treated as a background condition shaping content visibility, an emergent outcome shaped by platform conditions and user practices (Bishop 2019, 2589–2592). Under these conditions, creators operate in algorithmic opacity, inferring distribution logic indirectly while maintaining audience relationships and performing authenticity (Stewart 2018; Bonini & Gandini 2019; Frenneaux 2023). The second assumption concerns how observable outcomes are interpreted. Following the sociomaterial perspective (Orlikowski & Scott 2008, 434–455), outcomes on digital platforms emerge from the entanglement of technological infrastructures, platform conventions, and human practices—not from the actions of any single actor. This perspective informs the treatment of the unit of analysis: each TikTok post is approached as a post-assemblage through which these elements become empirically observable within the *S–St–R* framework (Bishop 2019; Gibbs et al. 2015, 257–259).

The first research question asks what content features and communicative messages are present in independent artists' high-engagement TikTok short videos. Within the *S–St–R* framework, stimulus (S) provides the analytical layer for addressing it: the immediately perceptible visual, sonic, textual,

and paratextual cues through which a TikTok video becomes classifiable and interpretable at first encounter, before any narrative development unfolds. This surface is analytically divided into content features and content messages—the two dimensions developed in Sections 3.2.1 and 3.2.2.

Content features are the observable formal and audiovisual properties that render a post legible in a fast-scrolling, context-collapsed feed, including video length, visual composition, interpersonal framing, platform-specific aesthetic conventions, metadata, and audio. These features serve as classificatory cues that enable viewers to rapidly infer the kind of content they are encountering and how to orient themselves toward the on-screen persona. Content messages, in turn, denote the immediate meanings that viewers derive from these cues at first exposure—meanings that operate at the level of initial interpretation rather than unfolding over time, providing a preliminary sense of what the video is about and how it is framed. In this study, content messages are captured through three dimensions—topic, vernacular genre, and affect—which together enable an unfamiliar viewer to orient to a post without prior knowledge of the creator.

The second research question asks how independent artists structure narrative meaning in these high-engagement videos through *micro-storytelling sequences (hook–progression–payoff)*. Staging (St) is the analytical layer through which this question is addressed: the temporal organization of stimulus-level cues within short-form videos—the ways creators sequence hooks, progressions, and payoffs under tight attention constraints so that viewers can recognize, follow, and resolve meaning almost instantaneously (Mueller & Rajaram 2022; Bouvier & Rasmussen 2022; Bösch & Divon 2024). Where stimulus describes what is in the post, staging describes what happens across it. The analytical focus is on how cues are aligned across time and modality to form recognizable narrative structures, marking the transition from meaning potential to meaning actualization as temporal alignment and rhythmic organization transform compositional cues into coherent narrative, affective, or evidentiary sequences.

The third research question asks what recurring *strategic archetypes* of artist-brand communication emerge across independent artists' high-engagement TikTok short videos. Response (R) completes the framework as the layer through which this question becomes empirically tractable. It is operationalized as the platform-recorded interaction signals embedded within each post—specifically comments, shares, and saves—treated as traces of how a post is taken up in practice rather than as measures of meaning in themselves (Dolan et al. 2016, 264–266; Muntinga et al. 2011, 15–16; Malthouse et al. 2013, 272). Each signal type reflects a distinct participatory orientation: comments index discursive engagement and affective reaction; shares indicate

circulation beyond the original viewing context; and saves signal the attribution of durable personal value. Together, these indicators provide the empirical basis for understanding how particular stimuli and staging configurations resonate within platform conditions.

Configuration-based research conceptualizes social phenomena as interdependent combinations of co-occurring elements, foregrounding relational coherence over isolated variables (Meyer et al. 1993; Fiss 2011). The present framework draws on this logic as a descriptive heuristic for identifying patterned strategic responses to platform conditions—arrangements in which content cues, staging practices, technological infrastructures, and user practices work together to produce visibility and participation on TikTok (Orlikowski & Scott 2008, 433–435; Bishop 2019; Martins et al. 2020; Curmi 2022; Jung 2024). Configuration-level patterns are traced only after stimulus and staging have been separated into their respective analytical layers, so that the relationship between them remains an empirical finding rather than an assumption built into the framework.

Configuration fluency names what is at stake when these patterns hold together. It describes a post-level capacity to assemble stimulus, staging, and response cues into arrangements that are simultaneously coherent to viewers and legible to the platform’s distribution logic. Empirically, this fluency is identified where compositional, temporal, and participatory dimensions align in ways that sustain attention and enable meaningful interaction within the constraints of short-form video. When configurations demonstrate internal coherence, recur across multiple posts, and remain empirically distinguishable, they can be differentiated as distinct pattern types and further developed into *strategic archetypes* in the empirical analysis—platform-native patterns through which independent artists negotiate the *paradox of independence*: the dual demands of algorithmic visibility and relational credibility (Gibbs et al. 2015, 257–259; Jung 2024).

Although stimulus, staging, and response are presented as analytically distinct layers, their sequential ordering is an analytical convenience: within the post-assemblage, they operate simultaneously and non-linearly, not as stages of production or experience. What the framework makes possible is a disciplined account of how independent artists assemble meaning under platform conditions—one specific enough to be empirically tractable and flexible enough to accommodate the variation that short-form content, by its nature, produces. Figure 3 illustrates the overall analytical framework. Chapter 4 sets out how that framework is put into practice.

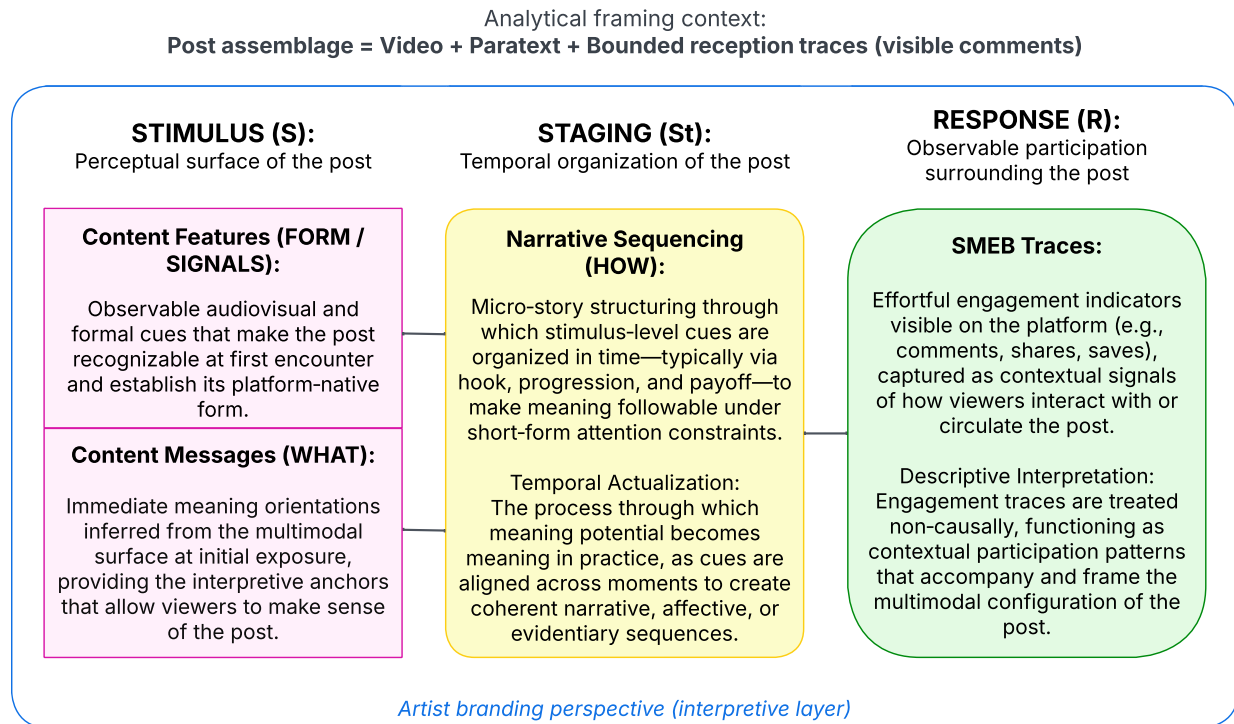


Figure 3. Analytical Framework for Examining TikTok Post-Assemblages (S–St–R Heuristic)

4 Methodology

4.1 Philosophical Approach

Deciding how to study something is never philosophically neutral. The choice of ontological and epistemological starting points shapes what counts as evidence, what kinds of claims the study can make, and how its analytical scope is defined (Kakkuri-Knuuttila 2006, 131–132; Eriksson & Kovalainen 2008, 13; Moon & Blackman 2014). In this research, interpretivist and social constructionist perspectives are combined to examine how communicative meaning is produced and made visible on TikTok. The implication is significant: such significance is not something content contains. It is something viewers produce. From an interpretivist perspective, it is understood to emerge through individuals' interpretations and actions, residing in the process of sense-making rather than inhering as a fixed property of content (Packard 2017; Pulla & Carter 2018; Ryan 2018, 8–10). Social constructionism further holds that significance is constructed and negotiated through social interaction, language, and shared practices—produced through relational engagement with media artifacts rather than deposited within them (Burr 1995, 2–3).

These perspectives jointly imply that social reality is understood as emerging from subjective interpretations (Eriksson & Kovalainen 2008, 15–18; Moon & Blackman 2014). In this thesis, 'success' is not assumed to be a stable or universal category but is approached as context-dependent and variably constructed across situations. Consistent with this orientation, the analysis focuses on the post-semblage as the empirical unit through which multimodal cues are examined, treating stimulus, message, and staging as analytical lenses for organising multimodal evidence. Artist branding is approached as a communicative practice that becomes analytically accessible through multimodal cues and platform-visible engagement traces.

One limitation raised against interpretivist and constructionist approaches is that they can foreground human meaning-making while underplaying the structuring role of technology, treating it as a passive backdrop (Winner 1993, 364–368; Mitev & Howcroft 2011, 300–301). To address this, the study adopts a sociomaterial perspective, recognizing the entanglement of social practices and technological infrastructures (Orlikowski & Scott 2008, 439–447). Accordingly, platform features such as interface design, audiovisual affordances, and recommendation logic are treated as active, constitutive conditions that shape how content becomes visible, circulates, and gains attention.

Given these commitments, abduction emerges as the most fitting mode of reasoning. Neither deduction nor induction, it is the logic of the best available explanation, revised as new evidence arrives. Classical formulations position it alongside deduction and induction as a distinct mode of inference (Peirce 1934, 117, 171, 181, 183; Peirce 1958, 217–218; discussed in Timmermans & Tavory 2012, 167–174), though its value here lies in its practical utility for interpreting multimodal, rapidly evolving platform content. Abduction involves iterative movement between empirical observations and theoretical concepts, with interpretations progressively refined in response to unexpected or resistant findings (Eriksson & Kovalainen 2008, 22–24; Timmermans & Tavory 2012, 170). In practice, abductive reasoning guided data familiarisation, memo-writing, and coding tool development as patterns stabilized—through revisiting coded segments, defamiliarizing close readings of short-form sequences, and weighing alternative theoretical framings until the most plausible interpretation held (Timmermans & Tavory 2012, 175–178; Eriksson & Kovalainen 2008). The resulting analytical categories reflect cumulative abductive insight rather than predetermined deductive templates. These four commitments—interpretivism, social constructionism, sociomateriality, and abduction—form a coherent philosophical stance that directly informs the research design.

4.2 Research Design

Understanding how independent artists design high-engagement TikTok content marketing strategies under algorithmic conditions requires a research design that is sensitive to patterns, context, and multimodal complexity. A qualitative, exploratory approach was adopted accordingly, allowing recurring strategies to emerge from the data without pre-specifying hypotheses and keeping category development anchored in analytically warranted observations. Because each post is a simultaneous assembly of visual, verbal, sonic, and gestural resources whose meaning arises from their interaction (Fazeli et al. 2023; Bouvier & Rasmussen 2022, 80), the analysis focuses on their configuration rather than on any single mode. The unit of analysis is the post-assemblage, as defined in Section 1.4.

The analytical framework is centered on *Visual–Verbal Video Analysis (VVVA)* (Fazeli et al. 2023). *VVVA* was selected because it translates multimodal analysis into a structured, six-step procedure detailed in Appendix 1 for segmenting, extracting, and coding visual, verbal, sonic, and gestural cues across temporally defined units—allowing narrative development to be traced as it unfolds within each post. Its preference over broader visual or discourse frameworks is both theoretical and practical: it couples a multimodality-grounded procedure with concrete extraction matrices suited to

short social-media videos, while delivering abductive flexibility and transparent cross-case comparability without requiring specialized software.

The analysis is theoretically informed by Multimodal Discourse Analysis, which underpins the understanding of meaning as emerging through the interaction of semiotic modes. The study also draws on principles of Qualitative Content Analysis to support systematic coding, category development, and cross-case comparison through data extraction matrices, thereby strengthening the transparency and consistency of the analytical process (Mayring 2014, 39–40; Schreier 2024, 15–19). The *VVA* coding procedure is paired with the *S–St–R* framework (Section 3.4), which serves as an organising logic for structuring the analytical process. This design makes possible a systematic account of how multimodal cues, temporal organization, and audience response work together within individual posts—the conditions necessary to answer all three research questions.

4.3 Data Sample and Unit of Analysis

Purposive sampling is the appropriate choice when the analytical objective is depth rather than breadth—when the goal is to understand how something works, not how widely it is distributed (Eriksson & Kovalainen 2008, 56–59). Qualitative research seeks analytical generalization and case-to-case transferability, not broad population-level deductions (Bouncken et al. 2026, 582). Accordingly, the analysis focused on high-engagement videos, as these cases provide clear, concentrated instances of visibility-building strategies, making the underlying mechanisms more conceptually accessible (Bouvier & Rasmussen 2022, 48).

The use of a focused sample is theoretically justified by the pursuit of in-depth contextual insight (Boddy 2016, 430) and by the principle of information power (Malterud et al. 2016, 1753–1756), which posits that a narrow research aim, high sample specificity, and a theory-informed analytical design together reduce the number of cases required while strengthening interpretive depth. In this study, these conditions are met through a tightly delimited research focus, a theoretically grounded analytical framework, and a detailed, video-by-video examination of multimodal content within a single-platform context, thereby supporting the adequacy of a relatively small but information-rich sample.

The final sample comprises 10 independent artists with an even gender balance (5 women and 5 men). For each artist, the three most engaging TikTok videos published between January and June 2025 were selected directly from the platform at the time of data collection, resulting in a corpus of 30 videos. Retaining several videos per creator was a deliberate choice: it enables observation of

how individual creators vary their communication strategies across posts while still permitting cross-creator comparison (Boddy 2016, 428–430; Malterud et al. 2016, 1753–1756). Gender parity (5/5) functions solely as a diversity safeguard; it is not analyzed. Sample size was determined ex ante to balance analytical depth and pattern detection: fewer than ten creators would risk insufficient recurrence of key patterns, whereas substantially more would reduce the depth of within-case analysis without advancing the study’s non-statistical objective. Given the study’s high sample specificity and theory-informed design, this configuration aligns with empirical methodological research indicating that qualitative saturation can be achieved with a relatively small, homogeneous sample of 9–17 cases (Hennink & Kaiser 2022, 114523), with core patterns often emerging within the first 6–7 units (Guest et al. 2020, 13). Three posts per creator thus provide a basis for systematic observation of within-artist variation, while a total of ten creators offers sufficient diversity for patterns to repeat across multiple artists. Sample adequacy was reviewed iteratively during coding (Malterud et al. 2016, 1756–1758).

For Research Question 3, the focus shifts from sample composition to the operationalization of recurring patterns. A configuration was treated as a recurring strategic archetype only when the same multimodal *S–St–R* pattern appeared across at least two to three different artists and in more than one post per artist. This recurrence criterion ensured that archetypes reflected robust patterns across multiple creators and posts, distinguishable from single-case anomalies. Outlier configurations were documented during coding but were not elevated to archetype status unless they met this predefined threshold. This threshold provides a defensible basis for claiming that a pattern is genuinely recurring—and not merely a feature of a single creator’s practice.

Creators were identified using Chartmetric to locate consistently active independent artists on TikTok; detailed discovery filters, the independence (®/©) criterion, and exclusion rules are documented in Appendix 2, and the anonymized case index and metadata schema in Appendix 3. Independence was operationalized from Spotify rights metadata (®/© lines) at the time of review, which form part of the structured metadata attached to digital music releases and distributed across platforms (Benham 2025; Spotify for Artists 2025); the ® symbol denotes the copyright holder of the sound recording, while the © refers to the underlying musical work (Copyright Service n.d.). When both lines named the artist or a non-major entity, this constituted a publicly verifiable signal of artist-controlled release practices. This operationalization enables consistent, transparent identification of independence at the level of observable release metadata. It is a principled and verifiable starting point—one that anchors the sample in publicly accessible evidence without overclaiming about the full complexity of artists’ industry arrangements.

The sample was further diversified across genres and broad geographic backgrounds, and all artists spoke and made music in English; however, these characteristics were not treated as analytical variables. For privacy reasons, creators are anonymized as Female Artist 1–5 and Male Artist 1–5. Additional background information and post identifiers are available under controlled access in Appendix 3. Each TikTok post is analyzed as a post-assemblage encompassing the video itself, its paratextual framing elements (captions, hashtags, mentions, on-screen text), and the interaction signals visible on the platform interface. These elements are treated as inseparable in the analysis, as paratextual cues shape interpretive framing (Gibbs et al. 2021, 5), while audience engagement serves as platform-mediated traces rather than direct indicators of internal states due to their algorithmic ordering (Bishop 2019, 2590–2592). Within this analytical structure, *micro-storytelling sequences (hook, progression, payoff)* function solely as secondary units that organize temporal development within the post-assemblage, without standing as independent units of analysis. This hierarchy reflects the multimodal and sociotechnical nature of TikTok content, in which meaning arises from the alignment of audiovisual, textual, and paratextual cues across the entire post, rendering isolated segments insufficient for interpretation.

All coding was conducted by a single coder throughout the study, ensuring interpretive consistency across cases while maintaining a continuous audit trail of segmentation decisions. Where finer temporal resolution was needed—particularly when tracing storytelling structures—the coder segmented videos into brief sequences or one-second intervals. Segmentation choices and coding decisions were recorded in the *VVA* instrument and associated matrices to preserve traceability (see Appendices 1 and 4). Engagement metrics were captured as time-specific snapshots of views, likes, comments, shares, and saves at the time of collection. To enable meaningful comparison across videos with different reach, engagement figures were standardized relative to views; the upload date was recorded to contextualize accumulation over time. Engagement is treated descriptively rather than causally, and its operational boundaries are discussed in Section 4.5.

Although the material consists exclusively of publicly available TikTok content, measures were taken to minimize the risk of identifying individual creators or users. Identities were anonymized; direct comment quotations were avoided; and distinctive wordings were paraphrased. No screenshots are included. These precautions reflect a responsibility to ensure anonymity and to handle research material with due ethical care, even when the data are public (Eriksson & Kovalainen 2008, 65–68). Interpretations remain traceable through the *VVA* matrices and supporting documentation (Appendix 3), and additional data handling safeguards are outlined in Appendix 5. Collectively, these decisions yield a dataset in which short-form communicative

strategies are observable and comparable across cases—the conditions necessary for a credible multimodal analysis of how independent artists design content for algorithmic visibility.

4.4 Visual-Verbal Video Analysis (VVVA)

Analyzing an independent artist’s TikTok post as a multimodal artifact requires an instrument capable of tracing visual, verbal, sonic, and temporal cues systematically across cases—and of doing so in a way that preserves the post’s internal coherence. *Visual-Verbal Video Analysis (VVVA)*, developed by Fazeli et al. (2023) and integrated here with the *S-St-R* framework, is that instrument. Two analytical boundaries govern the *VVVA* framework throughout: each observation is coded solely based on visual and auditory cues present in the post-assembly, and each is assigned to a single primary matrix (the single-location rule). A complete description of the *VVVA* procedure—including segmentation rules, multimodal transcription conventions, and hook/payoff boundary criteria—is provided in Appendix 1. The precise operationalization of each analytical dimension is detailed in Section 4.5.

To align the instrument with TikTok’s short-form logic, the coding tool is organized into four matrices: Matrix 1 catalogues stimulus-level form (e.g., audiovisual surface and on-screen paratexts), Matrix 2 traces *hook-progression-payoff* staging and audiovisual relations, Matrix 3 synthesizes message-level meanings (topic, vernacular genre, value orientation, affect), and Matrix 4 records platform-visible engagement indicators (comments, shares, saves). *VVVA* provides the multimodal evidence base, *S-St-R* structures the organization of observations, and their integration enables a systematic progression from close observation to cross-case interpretation. The matrix architecture is structured procedurally precisely because its function is to maintain analytical transparency when working with dense, multimodal material. The matrices serve as scaffolding for interpretive analysis—making the situated, theory-informed judgments central to this study traceable and comparable across cases.

4.5 Operationalization of Analytical Dimensions

Table 1. Operationalization of Analytical Dimensions

Research Question	Theoretical Concept	Analytical Category	Operational Definition	Visual/Multimodal Indicators	Sources
RQ1: What content features and content messages are present in independent artists' high-engagement TikTok short videos?	Platform vernacular; algorithmic legibility; multimodal affordances	Content Features (stimulus-level form cues; no interpretive meaning or affect)	Surface-level, visually and sonically observable cues that structure first-encounter readability and platform-native recognition; no meaning assignment or affect coding in this matrix.	Video duration and format; shot scale, camera angle and POV; lighting and visual composition; production aesthetic (raw vs. polished); on-screen text; captions, hashtags, mentions; audio source; editing rhythm	Gibbs et al. 2015; Han et al. 2020; Kress & van Leeuwen 2021
	Multimodal meaning-making; message framing; affect; artist branding	Content Messages (stimulus-level interpretive payload; includes affect)	The thematic, genre-based, value-oriented, and affective meaning that viewers derive from multimodal cues at initial exposure, including artist-branding relevant signals such as persona, relational positioning, and authenticity performance.	Topic/theme; vernacular genre; value orientation (informational, emotional, entertainment, transformational); dominant affect (valence–arousal); performer stance	Han et al. 2020; Russell et al. 1980; Schreiner et al. 2021; Grzenkiewicz & Wildfeuer 2025; Fazeli et al. 2023
RQ2: How do independent artists structure narrative meaning in high-engagement TikTok short videos through micro-storytelling sequences?	Microstorytelling ; narrative compression; temporal sequencing	Staging (hook–progression–payoff organization; temporal sequencing only)	Observable temporal sequencing through which multimodal cues establish an hook, guide interpretation through progression, and deliver a payoff that resolves or fulfils viewer expectations under short-form constraints; no value or affect assignment is conducted in this matrix.	Hook window (0–3s); hook cues (visual, textual, sonic); sequencing of frames, gestures, edits, and sound; escalation or modulation; payoff moment (musical, affective, narrative, or informational); payoff onset; audio–visual relation (aligned/complementary/divergent)	Mueller & Rajaram 2022; Bouvier & Rasmussen 2022; Fazeli et al. 2023
RQ3: What recurring strategic archetypes emerge across successful independent artists' high-engagement TikTok short videos?	Configuration logic; sociomaterial meaning-making; SMEB (Social Media Engagement Behaviors)	Strategic Configuration (cross-layer S–St–R synthesis; non-causal)	Recurring co-occurrence patterns linking content-feature cues, message-level meanings, staging devices, and platform-visible audience participation traces into stable strategic archetypes, interpreted through configuration-level saturation rather than variable-based causal explanation.	Cross-matrix patterning of features, messages, staging sequences, and engagement traces	Meyer et al. 1993; Fiss 2011; Dolan et al. 2016; Muntinga et al. 2011

	Platform participation; engagement behavior	Response (platform-visible engagement traces; descriptive, non-causal)	Observable interaction signals embedded in the post-assemblage, including comments, shares, and saves, treated as distinct participatory orientations (discursive engagement, circulation, and value attribution) that contextualized how content configurations are taken up by audiences. These signals function diagnostically as empirical traces rather than causal outcomes.	Comments, shares, saves; per-view ratios; engagement mode; reception climate	Dolan et al. 2016; Muntinga et al. 2011; Malthouse et al. 2013
--	---	--	--	--	--

Translating the analytical framework into practice demands careful calibration. The core constructs—stimulus, staging, and response—must be rendered observable in ways that preserve the multimodal and temporally unfolding character of short-form TikTok content. This study treats operationalization as an interpretive commitment: a set of principled decisions about which aspects of the post-assemblage can be made analytically tractable, under what conditions, and on what grounds. This orientation follows from the abductive and sociomaterial foundations outlined earlier in this chapter and aligns with the broader turn in multimodal research toward treating analytical categories as relational rather than discrete. Within this logic, the *S–St–R* dimensions are treated as analytically distinct yet co-present layers—each becoming observable through different, but coordinated, registers of the post-assemblage.

The conceptual operationalization used in this study is summarised in Table 1, which situates the analytical categories within the broader theoretical framework, shows how each research question maps onto a specific analytical layer, and functions as an orienting device making the analytical architecture visible to the reader at a glance; the full coding specification is documented in Appendix 4.

Building on this mapping, the analysis specifies how each layer is operationalized in practice. At the stimulus layer, cues are identified by distinguishing between formal features—those that structure how the post is immediately perceptible—and meaning-bearing elements that become legible only as the sequence unfolds. This distinction is necessary because identical visual or sonic cues may serve different interpretive functions depending on their temporal placement and co-occurrence with other elements; collapsing form and meaning at this stage would reduce analytical precision. Staging is examined through observable cue sequencing, with the coder tracing shifts in rhythmic density, modal convergence, and transition structure to document what the arrangement of cues makes available to a viewer over time. Response is captured through platform-visible engagement data, which provides a partial, already filtered trace of circulation; this condition renders a descriptive rather than causal treatment analytically appropriate. Normalized per view and assessed through mode-dominance thresholds, engagement indicators enable comparison across posts of varying reach, with the response layer functioning as an interpretively contextualizing presence that enriches the stimulus and staging analyses.

These analytical commitments can be illustrated through a simplified example. Consider a post in which an independent artist performs a short vocal excerpt recorded in a single, continuous shot with synchronized subtitle text. At the stimulus layer, the observable features consist of surface-

level cues such as static framing and real-time lyric subtitles. At the staging layer, the post follows a continuous temporal progression without cuts, with minimal visual variation aligned with the vocal delivery. At the response layer, the post yields platform-visible participation traces, normalized per view. The example demonstrates how the same artifact becomes analytically visible in different ways depending on whether attention is directed to form, temporal organization, or circulation traces. This operationalization grounds each research question in observable data while maintaining analytical differentiation across layers. How that framework was itself developed is the subject of the next section.

4.6 Abductive Development of the Coding Framework

The coding categories were defined in advance to support transparency and comparability, while the framework itself was designed to remain responsive rather than fixed. Its final form emerged abductively as the researcher worked through the data, with early assumptions refined in light of what the material revealed. During pilot coding, several theoretically derived distinctions required adjustment to better capture the multimodal and temporally compressed nature of short-form TikTok videos. Some staging distinctions, initially drawn from prior literature, were revisited when the researcher encountered posts in which narrative movement and sensory payoff merged into a single uninterrupted sequence—as happened, for instance, when a trending audio cue simultaneously introduced and resolved the informational premise within the first two seconds. Certain message-level categories were also re-specified to accommodate hybrid informational–musical resolutions that the theoretical model had not initially anticipated. These refinements were not departures from the framework so much as developments that made it more responsive to how meaning materialized in the material itself.

As the analytical scheme evolved through this abductive process, it became clear that the most productive distinctions were temporal. What mattered most was the speed and manner in which cues aligned, diverged, or resolved over time. This shifted the analytical emphasis away from enumerating discrete elements and toward organising them so that patterned regularities could emerge across cases. The framework thus operates as a configurational device in which patterns arise from the interplay among stimulus-level features, temporal sequencing, and platform-visible traces.

The matrix structure outlined in Section 4.4 proved useful for operationalizing this orientation, as separating the three layers enabled systematic cross-case comparison without compromising the integrity of each post-assembly. These layers are analytical rather than ontological distinctions;

their value lies in providing a consistent scaffold for tracing recurring temporal multimodal content design. Maintaining it across cases required sustained analytic discipline. It required holding a clear boundary between descriptive and inferential statements—a distinction that demands careful attention when working with TikTok material. The single-location rule established earlier was important here. It functioned as more than a procedural safeguard—it represented an epistemic commitment: each interpretive move had to be grounded in a specific empirical anchor. This kind of traceability matters in any single-coder design. It becomes especially pressing in cross-modal contexts where several plausible readings can coexist without any one of them being obviously wrong. Restricting each observation to a single primary location within the coding architecture served two related purposes: it kept analytic drift in check and prevented any single cue from accumulating interpretive weight across multiple categories.

The framework was built for a specific analytical task—cross-modal analysis of short-form TikTok posts within a theoretically informed qualitative sample—and it is precisely that specificity that makes its findings interpretively defensible. It provides a transparent and auditable way of organising semiotic evidence while preserving the relational complexity through which meaning emerges. Full technical documentation is provided in Appendix 4.

4.7 Trustworthiness of the Study

In a single-coder, multimodal study where interpretive judgments cannot always be reduced to explicit decision rules, the question of trustworthiness is not procedural formality but a genuine analytical obligation. It is assessed here using four criteria: credibility, transferability, dependability, and confirmability (Eriksson & Kovalainen 2008, 308). These principles address how interpretations are grounded in the data, how openly the research process is documented, and the extent to which findings carry analytical purchase beyond the immediate dataset.

Credibility refers to the extent to which interpretations are grounded in the data (Eriksson & Kovalainen 2008, 306). Here, it is built through consistent cross-referencing between coded multimodal observations and the patterns that emerged across posts. The structured coding frame—described in Section 4.6—enabled systematic tracing of *stimulus–staging–response* configurations. At the same time, sustained engagement with the material allowed theoretical concepts to develop from the data rather than be imposed on it. (Corbin & Strauss 2012, 24.) Repeated close readings were part of this process, and where early interpretations were revised in light of additional cues or competing explanations, those revisions were recorded—making the development of the analysis itself part of the credibility record.

Transferability addresses whether findings carry meaning in other contexts (Eriksson & Kovalainen 2008, 308–309, 331). To support this, the researcher has provided detailed accounts of the platform context, the sampling rationale, and the analytical procedures, giving readers sufficient grounding to assess the approach's applicability to comparable digital settings. The findings reflect context-bound patterns observed in short-form music-related content on a single platform over a defined period, and their value for other researchers lies in the analytical logic they demonstrate—one that is transferable to other short-form platforms or content categories where multimodal staging and audience engagement intersect.

Dependability relates to the consistency and stability of the research process (Eriksson & Kovalainen 2008, 308). In qualitative work, this requires that the reasoning behind each analytical decision is recoverable and auditable, even as the analysis remains responsive to the material. The matrix-based coding structure provided a systematic backbone, while the audit trail ensured that departures from initial categories were recorded rather than silently absorbed into the analysis. As a further check, a random subset of five posts was recoded after a two-week interval; the consistency of results gave the researcher confidence that the coding frame was being applied consistently throughout the analytical process. Analytic memos and decisional records captured how ambiguous cases were worked through and how competing readings were evaluated during the abductive iteration process detailed in Section 4.6 (see Appendices 4 and 5). Engagement figures are time-stamped throughout, preserving temporal comparability and keeping the analytical record auditable from start to finish.

Confirmability concerns whether findings reflect the empirical material rather than the researcher's own assumptions or preferences (Eriksson & Kovalainen 2008, 308, 324). This was perhaps the most demanding criterion to address, given that multimodal analysis involves interpretive judgments that cannot always be reduced to explicit decision rules. Transparent coding guidelines and systematic cross-referencing between descriptive observations and interpretive claims were central to the approach. Boundary cases were documented rather than quietly resolved, and artifact-level analysis remained the primary focus throughout—audience engagement signals were drawn on to contextualize meaning, without being treated as definitive. Interpretive claims were accepted where patterns held consistently across multiple layers of the *VVA* matrices, a requirement that guarded against undue weight being placed on any single cue. Supporting materials are included in Appendix 4.

Analyzing short-form multimodal content posed its own interpretive challenges—ones that no protocol fully eliminates. Meaning in this kind of material does not always surface in obvious ways—it often emerges through subtle alignments between elements that resist neat categorization. Abductive reasoning meant sitting with uncertainty at times, returning to borderline cases more than once to test whether the data could justify the preferred reading. The researcher recognizes that subjectivity is a constitutive feature of qualitative work, one to be acknowledged and managed rather than treated as an external contaminant. What the documented reasoning process, memo-writing, and cross-layer verification are intended to provide is accountability: a record of how each interpretive step was reached and the evidence on which it rests.

4.8 Ethical Considerations

Platform-based research involving publicly accessible digital traces carries ethical obligations that public accessibility alone does not dissolve. This study is conducted in accordance with the European Code of Conduct for Research Integrity (ALLEA 2023), the GDPR, and the University of Turku's ethical guidelines—commitments that translate, in practice, into transparent analytical documentation, careful handling of digital traces, accurate representation of data, and clearly justified data-collection boundaries. Privacy, contextual integrity, and the potential for harm remain live concerns even when the material is publicly posted. Technical details on AI tool settings and data-sharing controls are provided in Appendix 6, and data governance, storage, retention, and version control practices are documented in Appendix 5.

Anonymization was implemented with particular care due to the re-identification risks associated with platform-based content. Public social media traces are easily searchable, persistent, and replicable (boyd 2011, 61), and contextual details can enable identification even after direct identifiers are removed (Zimmer 2010, 313; Hoffmann & Jonas 2016; Ingierd & Fossheim 2016). In response, all creators are anonymized, distinctive textual elements are paraphrased, and screenshots are not included. Structured supporting materials—such as the controlled-access anonymized case index (Appendix 3) and composite coding examples/codebook documentation (Appendix 4)—maintain transparency while reducing identifiability.

A further consideration arises from the platform's own architecture. Ranking mechanisms determine which comments become visible, so top comments reflect the algorithmic ordering within a filtered selection of audience perspectives (Bishop 2019, 2590–2592). Consequently, as noted in Section 1.4, audience engagement signals are treated as contextual traces rather than representative opinion data, with interpretation anchored in the material directly visible within the

post-semblage. The analytical scope is defined accordingly: claims are grounded in publicly observable evidence, and the study addresses communicative strategies and multimodal content design rather than audience psychology or the internal workings of platform recommendation systems. These practices reflect a broader commitment to handling platform-based material responsibly and to keeping interpretive claims within the evidential boundaries the data can sustain.

5 Findings

This chapter presents the study's empirical findings. The question driving the analysis was, in that sense, a practical one: how do artists who lack institutional support achieve visibility with audiences who have no prior reason to stop scrolling? What the data reveal is a notable degree of patterned regularity—a repertoire of multimodal content design through which posts achieve legibility under interest-graph distribution—a finding that holds across artists who differ markedly in genre, aesthetics, and audience scale. That regularity is the chapter's central subject.

The recurring multimodal configurations identified here are treated as descriptive patterns that organize observable regularities within post-assemblages and reveal structural consistencies across high-engagement posts. The analysis proceeds through three interrelated layers derived from the *S–St–R* framework: the stimulus-level features that render each post perceptible at first encounter; the *micro-storytelling* structures through which meaning is temporally organized; and the configuration-level combinations that connect stimulus, staging, and platform-visible traces into *strategic archetypes*. Audience-engagement indicators are read as contextual signals that accompany the multimodal construction of meaning without determining it—the higher-order configurations synthesized in Section 5.3 are drawn from precisely this relationship. Section 5.4 then gathers these findings, connects them to the literature reviewed in Chapters 2 and 3, and develops the enriched theoretical framework that the study contributes; detailed row-level coding evidence and engagement-mode calculations supporting Research Questions 1 and 2 are documented in Appendix 7.

5.1 Content Features and Communicative Messages in High-Engagement TikTok Videos

What makes a TikTok post immediately legible to a viewer who has never encountered the artist before? Research Question 1 addresses this by distinguishing two analytically separable but co-present dimensions of stimulus-level meaning. Content features—operationalized in Matrix 1—are the visually and sonically observable formal cues through which a post becomes perceptible at first encounter. Content messages—operationalized in Matrix 3—are the meaning payloads that can be inferred from the post's multimodal composition once its formal surface has been registered, encompassing topical focus, vernacular genre, communicative value orientation, and affective stance. Both dimensions belong to the stimulus layer within the *S–St–R* framework, though they operate at different levels of abstraction: features establish the perceptual conditions for legibility;

messages specify what that legibility communicates. Subsection 5.1.1 reports the aggregated stimulus-level features found across all 30 posts; subsection 5.1.2 turns to the message-level meaning patterns through which these posts become interpretable within TikTok's short-form context.

5.1.1 Content Features and Multimodal Form

The 30 posts share no single aesthetic formula. Production registers split roughly evenly between Raw/Lo-Fi and Polished/Hi-Fi presentations without genre-specific clustering. Yet the posts are unmistakably legible as belonging to the same platform. That is the first finding worth pausing on. What makes a post recognizable, then, is not its production quality but the coherence of its overall multimodal configuration. This shared perceptual infrastructure operates independently of any individual creator's aesthetic identity and ensures that posts are immediately classifiable to unfamiliar viewers within TikTok's interest graph. All posts employ the standard TikTok format; none use Duet or Stitch compositions. Screen recording appears in only a small minority of cases, confined to process-demonstration contexts, and video durations remain within platform-typical short-form boundaries throughout.

The choice of setting is the first relational cue a viewer registers, and the corpus indicates that it operates with considerable consistency. Five spatial configurations recur: domestic interiors, street locations, nature-based exteriors, studio or performance spaces, and car interiors. Domestic interiors are the most common, typically featuring unstaged rooms with stable, naturalistic lighting. Street contexts introduce ambient movement and public-space cues; nature-based settings open into wide visual fields; studio and performance spaces are distinguished by controlled lighting and branded or professional elements. Car interiors appear least frequently but form a stable minor category shaped by proxemic constraints. The variety across these settings is less analytically significant than what each setting communicates before any explicit message is delivered. Domestic and car interiors carry an implicit register of intimacy and access—positioning the artist as approachable and proximate—while studio and street settings tend to signal professionalism or spontaneity. Setting choice, therefore, functions as an early-stage cue that contributes to perceived relational proximity—the sense of intimacy between audience member and media figure (Sood & Rogers 2000, 387; Stewart 2018, 56)—positioning the artist as either intimate and accessible or performatively distant.

Framing choices follow a similarly constrained logic. Close-up shots dominate in performance-oriented home recordings and screen-recorded hybrids; medium framing characterizes studio-based

or narrative sequences; long-shot framing is confined to a small subset of outdoor or multi-set clips. The most pronounced pattern concerns camera angles: nearly all posts adopt an eye-level view, with overhead or steep-angle shots appearing only when a specific artifact needs to be displayed. Eye-level framing establishes a default condition of perceptual equality between artist and viewer—prior research demonstrates that visual proximity, such as close-up presentation, can enhance perceived intimacy by creating an illusion of face-to-face interaction (Sood & Rogers 2000, 387, 389). Point of view is more variable: direct second-person address recurs across performance and diary formats; observational third-person vantage points characterize narrative and reaction clips; first-person perspectives surface primarily in screen-recorded content. Across all three modes, the underlying logic is relational. Framing positions the viewer within the post's communicative frame before any explicit message is registered.

On-screen text is used selectively and with evident intentionality. Subtitles appear in more than half of the posts—most consistently in vocal-performance contexts, where time-aligned lyric transcription supports clarity without imposing additional semantic framing. Overlay text delivers brief situational cues in fewer cases, and several posts dispense with text entirely, relying exclusively on audiovisual composition. What is worth noting is that individual creators tend to maintain a characteristic text-use strategy across their posts—a pattern that positions text deployment as part of each artist's broader legibility repertoire, one that persists across varied content types and therefore carries analytical weight beyond any single post.

Paratextual metadata indicates a bimodal pattern: most posts feature very short captions and few hashtags, while a smaller subset uses high-density metadata with genre- or trend-related tags. Mentions are sparse and, where they appear, reference brands, collaborators, or fellow artists. These paratextual elements function as surface-level discoverability signals—operating at the platform's categorical layer, distinct from the multimodal meaning construction that unfolds within the post itself.

Audio sourcing splits between TikTok library tracks and original sound, with original audio particularly common in creator-produced or process-oriented posts. Library audio appears across both polished and lo-fi registers—a finding that complicates any straightforward mapping of original versus borrowed sound onto production quality. More consequential than sourcing is timing: audio consistently enters within the opening seconds, and in many posts, vocal delivery, gestures, or movement is synchronized with the track in ways that directly contribute to the hook's multimodal alignment. Audio-ownership labels confirm that many creators use sound attributed to

themselves—a paratextual detail that reinforces the authenticity signals constructed at the message layer.

Across the corpus, these patterns point toward a stable set of formal cues that enable rapid perceptual legibility under conditions of context collapse. Female Artist 1__1's marketside performance illustrates this precisely: the combination of eye-level framing, direct address, and immediate subtitle onset occurs within the first second, making the post's interpersonal stance legible before any semantic interpretation begins. What is most analytically significant is the consistency of this infrastructure across artists who otherwise differ markedly in genre, aesthetic, and audience size. It functions as a kind of shared vocabulary that unfamiliar viewers decode almost automatically. From this foundation, message-level meaning and staging patterns subsequently build.

5.1.2 Message-Level Meaning Patterns

Where the stimulus-level findings show how posts become perceptible, the Matrix 3 data reveal how they construct meaning—and the pattern is one of deliberate thematic constraint, with topical choices converging in ways that indicate an underlying structural logic beyond personal preference. Topical orientation clusters into three recurring groupings. The largest—a performance-and-promotion cluster—comprises six release-adjacent posts and four additional performance- or preview-led clips in which meaning is anchored in musical delivery, stylistic presence, or promotional signals. A relatability-and-interaction cluster draws together five everyday-life posts, three community-interaction cases, and one social-dynamics piece, each generating meaning through interpersonal familiarity or situational resonance. A process-and-craft cluster comprises four tutorial or how-to posts, one aesthetic world-building case, and one behind-the-scenes creative-process sequence, each foregrounding artistic technique as the primary source of meaning. Several isolated topics appear at the margins. These artists do not range widely—they return repeatedly to the same interpretive territory. That thematic precision is itself a legibility strategy.

Vernacular-genre tendencies reinforce this picture. Vlog- and diary-style formats dominate, appearing in 14 posts and consistently offering a direct, interpersonal interpretive frame. Performance genres account for 10 posts, with musical or choreographic display as the central expressive device. Tutorial and demonstration genres appear in four posts; storytelling or confessional formats surface less frequently but offer targeted narrative framing when they do. The consistency of genre clarity across the corpus is analytically significant. Every post maintains a

single, legible vernacular genre—a consistency that functions as a prerequisite for the rapid interpretability that TikTok’s interest-graph distribution demands.

Genre tells the viewer how to read a post; communicative value orientation tells them why it is worth reading. The distinction matters because two posts can share the same vernacular genre—both formatted as vlogs, for instance—and yet offer entirely different experiential payoffs. That difference becomes decisive under algorithmic conditions where the viewer’s decision to continue watching is made almost instantaneously. Communicative value orientations reveal a similarly concentrated pattern: entertainment value is primary in fourteen posts, manifesting through aesthetic vividness, energetic delivery, or humor; emotional value anchors eleven posts centered on vulnerability, relational tone, or lyrical expressiveness; informational value is confined to three procedural demonstration posts; transformational value to two reflective or identity-driven clips; and transactional value appears only once, embedded within a broader process-narrative. The more analytically significant pattern emerges in the secondary codings: entertainment and emotional value co-occur with notable frequency, functioning as mutually reinforcing registers within short-form multimodal expression. This combination accounts for the particular affective intensity characteristic of high-engagement independent artist content.

Affective orientation extends this picture in a direction that resists simple summary. High-arousal affect dominates, appearing across 22 posts and spanning positive, negative, and mixed valence states. Mixed valence is the single most common affective configuration—present in 16 posts—and tends to mark content characterized by emotional complexity or intensity rather than straightforward expressiveness. Positive valence appears in 11 posts through lighthearted or bright affect; negative valence is less common and typically associated with confrontation or tension. Neutral arousal is confined to eight posts, concentrated in reflective diary entries and instructional content, where a measured tone becomes part of the communicative stance itself. The prevalence of high-arousal across the corpus appears to serve a pragmatic legibility function—amplifying interpretive signals in a feed where attention is contested at every scroll—a reading that finds support in prior research showing that high-arousal content is more likely to be shared and to generate engagement (Berger & Milkman 2012; Nelson-Field et al. 2013, 210).

Payload drivers specify how meaning becomes actionable for the viewer. Relatability is by far the most frequent, appearing in 16 posts—a dominance that deserves direct attention. These artists are not primarily selling a product but constructing a relational position, and that distinction carries substantive consequences for how independent artist branding on TikTok should be understood.

Musical payoff anchors four performance-focused posts; call-to-action structures anchor four promotion-oriented clips; personal-story drivers and humor each appear in three. Supplementary cues—vulnerability in five posts, humor in six—further refine interpretive stance without displacing the primary driver. Reciprocity prompts are present in most posts but tend to remain implicit, woven into relational framing as an ambient invitation rather than an overt request.

These patterns cohere into several stable meaning configurations that, once identified, are difficult to miss across the corpus. An entertainment-high-arousal performance configuration links performance genres, entertainment value, and musical payoff. A relatable-emotional diary configuration aligns vlog framing with drivers of mixed or positive valence, high arousal, and relatability. An instructional-informational configuration integrates neutral arousal, procedural demonstration, and informational value. A less frequent transformational configuration embeds personal change within high-arousal expressive framing. Across all four, high-engagement meaning arises from stable, recognizable multimodal structures that remain interpretable under the constraints of context collapse. The stimulus- and message-level findings, read together, suggest that independent artists rely on a recurrent multimodal system that functions as something like an algorithmic recognizability grammar—a shared logic of legibility through which content must signal its nature, stance, and communicative intent within the first few seconds, or risk being discarded in the feed. This is the most analytically generative finding at the RQ1 level. The RQ1 findings thus provide the conceptual basis for the temporal sequencing mechanisms examined in Section 5.2, where staging transforms these recognizability structures into temporally organized *micro-storytelling* forms.

5.2 Micro-Storytelling Sequences in High-Engagement TikTok Videos

Table 2. Analytical Dimensions of Micro-Storytelling Staging Configurations (Immediate Loops, Revealed Progressions, and Escalation Arcs)

Analytical Dimension	Immediate Loops (Arc 0)	Revealed Progressions (Arc 1)	Escalation Arcs (Arc 2)
Dominant hook modality	Hooks begin through tightly aligned multimodal synchrony, with audio onset, gesture, gaze, and camera entry converging within the first fractions of a second.	Hooks typically emerge through aligned or visual-anchored cues—spoken frames, overlays, or subtle movement that preview an impending reveal or process.	Hooks are aligned but oriented toward establishing an initial premise designed to support subsequent escalation rather than immediate closure.
Typical hook logic	Promise/Outcome or immediate pattern-break cues that deliver meaning almost instantly.	Promise/Outcome or curiosity-driven cues that establish a premise for sequential revelation.	Curiosity- or surprise-based cues that intentionally open space for tension or contradiction.
Progression logic	Minimal or no progression; meaning stabilises immediately and unfolds through a continuous expressive plateau.	Procedural accumulation of micro-cues that build toward a reveal, demonstration, or clarifying event.	Presence of a complication phase involving rising tension, contradiction, or misdirection before payoff.
Payoff type	Almost exclusively musical or affective, delivered through instantaneous rhythm–gesture synchrony.	Reveal-based, musical, or CTA-integrated payoffs depending on the post’s premise and communicative aim.	Punchline- or conflict-resolution-based payoffs that depend on delayed closure.
Payoff timing	Immediate payoff, typically within the first moments of the clip.	Early-to-mid-window payoff resulting from sequential layering.	End-window payoff that resolves built-up tension or incongruity.
Audio–visual relationship	Strongly aligned across modalities, creating instantaneous interpretive coherence.	Primarily aligned, sometimes enriched with complementary visual layers such as contextual overlays or background cues.	Aligned or intentionally divergent, depending on how escalation and misdirection are used to shape tension.
Narrative function	Provides instant experiential closure; the narrative is the sustained multimodal moment itself.	Builds meaning procedurally, with each micro-cue contributing to the unfolding narrative direction.	Manipulates expectation through tension, inversion, or contrast before offering closure.
Interpretive affordance	Meaning is synchronous with performance from the outset and requires no temporal elaboration.	Meaning accumulates progressively and is retrospectively consolidated at the moment of payoff.	Meaning depends on the release of tension and is fully interpretable only after resolution.
Role within S–St–R logic	Staging-as-Immediate-Fulfillment.	Staging-as-Sequential-Revelation.	Staging-as-Tension-Modulation.

Posts of only a few seconds can enact fully intelligible narrative arcs. That finding is less obvious than it appears: it means that temporal organization, not duration, is what makes meaning cohere in

short-form video. The Matrix 2 data show this operating through *micro-storytelling*—a meaning-making mechanism in which multimodal cues are temporally organized to generate interpretive expectations, modulate tension, and deliver resolution (Mueller & Rajaram 2022, 25).

Interpreted abductively across the corpus, these staging practices consolidate into three stable temporal configurations—*Immediate Loops*, *Revealed Progressions*, and *Escalation Arcs*—whose defining analytical dimensions are summarised in Table 2. What these configurations reveal is how recognizability at the stimulus level (RQ1) becomes narratively operative only when temporally sequenced into an expectation–tension–resolution flow: *Immediate Loops* through instantaneous synchrony, *Revealed Progressions* through incremental clarification, and *Escalation Arcs* through tension and delayed closure. In all three cases, narrative meaning arises from the timing of multimodal signals relative to one another.

Across the sample, the hook window—typically observable within approximately the first one to three seconds—functions as the decisive moment in which interpretive expectations are established. Hooks simultaneously capture attention and frame the interpretive contract by signaling the post's genre, stance, and likely direction. The empirical pattern is unambiguous: when multimodal elements converge rapidly within the hook, viewers remain engaged; delayed alignment weakens the interpretive pathway. In high-engagement posts, hooks perform a dual epistemic function—orienting the viewer and simultaneously constraining the range of plausible readings. This is a finding in its own right, not merely a precondition for the configurations that follow.

Most posts employ aligned multimodal hooks in which audio onset, gesture, gaze orientation, and camera stabilization co-occur. In Female Artist 1__1's marketside performance, the beat drop coincides precisely with her step into the frame, her hip shift lands on the sonic accent, and the subtitle appears in synchrony with the opening lyric. This convergence frames the clip as a performance-driven sequence, allowing viewers to infer the communicative stance almost instantly. In other cases, sound leads before visual. In Male Artist 2__2's outdoor performance, the opening lyric arrives while the camera is still stabilising, producing an audio-led hook that guides interpretation before gesture and framing catch up. Visual-led hooks foreground overlays or salient visual contrasts, as in Female Artist 3__1's post, where the “POV you have auburn hair” text sets the interpretive premise that subsequent visual shifts reinforce. Across these types, hooks function by establishing the viewer's interpretive expectations—ones that the progression must then either confirm, elaborate on, or strategically delay.

Following this initial alignment, the progression stage manipulates informational momentum or tension, depending on the configuration. In *Immediate Loops*, hook and payoff merge into a single expressive moment: the interpretive expectation created in the hook is fulfilled almost at once, producing experiential rather than narrative closure. Meaning arises through experiential synchrony—an affective mechanism in which rhythmic coherence substitutes for narrative development. Female Artist 1__3's synchronized dance sequence exemplifies this precisely: the beat drop and first movement coincide, and the video sustains this unified kinetic pattern throughout, with no interpretive gap between what the hook promises and what the post delivers. That economy of expectation is itself the branding move—the artist appears as someone who gives the viewer exactly what they came for, immediately.

Revealed Progressions operate through a different mechanism, one in which progression functions as an epistemic arc—a sequence of micro-cues that gradually reduce interpretive uncertainty until closure becomes possible. Female Artist 3__2's vinyl-making post demonstrates this with particular clarity: each new visual element—materials floating on water, close-ups, shifting light—serves as accumulating evidence toward the anticipated reveal, and the final display of the vinyl artwork retrospectively organizes the preceding cues into a narrative whole. Male Artist 1__3's notebook reveal follows the same logic, where an initial calm spoken frame establishes the premise and the later disclosure of handwritten lyrics delivers resolution. The branding implication is distinct from that of *Immediate Loops*: where experiential synchrony positions the artist as emotionally present, sequential clarification positions them as credible and transparent—someone whose process is worth following through to the end. Meaning is procedurally earned.

Escalation Arcs introduce a complication between hook and payoff, with tension modulation as their core mechanism: they create interpretive disequilibrium that is resolved only at the payoff. In Female Artist 4__2's confrontation post, the hook sets the stage for conflict, after which intensifying gestures and emotional cues heighten uncertainty until a final pivot resolves the scene. Male Artist 3__1's institutional sketch employs misdirection—a mundane setup cues a familiar reading that is then inverted by a sudden recontextualization. The viewer must tolerate uncertainty until the payoff reorganizes earlier cues into a meaningful whole. When the tension–resolution sequence is timed precisely, the affective release is heightened—and it is that release, more than any individual cue, that makes these posts shareable. The branding position produced here is memorability: the artist who can make a viewer feel the turn earns a different kind of attention than one who simply performs.

Despite their structural differences, all three configurations depend on multimodal synchrony within the hook as the foundation of interpretive intelligibility—audio-visual alignment, rhythmic consonance, gaze coordination, and subtitle timing all contributing to the conditions under which viewers can immediately infer how to read the post. The hook is not an introduction. It is an argument. Divergent audio-visual relations succeed only in controlled comedic contexts, where intentional mismatch itself becomes the communicative device. Staging is a temporal strategy for legibility, not aesthetic embellishment. Posts must create coherence at the micro-temporal scale to remain interpretable under interest-graph distribution—and it is precisely that coherence, engineered across seconds, that separates high-engagement posts from those that lose the viewer before the payoff arrives.

Across all three configurations, TikTok *micro-storytelling* emerges as a configuration-level system of temporal recognizability: meaning arises from the timing of multimodal cues, not from any single element considered in isolation. *Immediate Loops* offer instantaneous experiential closure; *Revealed Progressions* construct meaning through sequential elaboration; *Escalation Arcs* generate it through tension and release. What the RQ2 findings add to the RQ1 findings is the temporal dimension—the demonstration that stimulus-level recognizability becomes narratively operative only when its cues are sequenced into an expectation–resolution flow. In *S–St–R* terms, staging is the organizational mechanism through which stimulus cues acquire temporal coherence and the viewer’s interpretive pathway is guided toward resolution. The next Section 5.3 shows how these staging mechanisms combine with stimulus- and message-level patterns to produce higher-order *strategic archetypes* of artist-brand communication—the level at which the framework’s explanatory reach is fullest.

5.3 Strategic Archetypes of Artist-Brand Communication in High-Engagement TikTok Videos

The third research question examines recurring *strategic archetypes* arising from the patterned co-occurrence of stimulus features, message meanings, staging practices, and platform-visible response traces in high-engagement posts. At this cross-layer level, the analysis moves beyond isolated multimodal cues to examine how content is encountered under conditions of context collapse—situations in which multiple, heterogeneous audiences converge and interpret it through divergent normative frameworks (see Palombo 2024, 95–105). The archetypes identified here represent replicable configuration strategies through which creators negotiate the demands of TikTok’s legibility. They are distinguished from idiosyncratic stylistic habits precisely by their recurrence

across multiple artists and multiple posts: what appears in only one creator's practice is not elevated to archetype status. This is also the point at which the artist-branding lens becomes most analytically productive: the archetypes not only organize multimodal construction but articulate recognizable branding positions that integrate expressive immediacy with platform-native distribution logic.

Four higher-order archetype families emerge across the corpus. Each is analytically distinct from conventional content genres, audience segments, or creator identities: they represent multimodal configuration solutions through which creators negotiate the dual pressures of TikTok's algorithmic visibility system and the relational expectations that define independent artist branding. Each archetype integrates a distinct combination of stimulus-level signals, staging logics, interpretive meanings, and response climates—reflecting how creators position themselves, through divergent multimodal routes, as credible, affective, or socially portable figures within the attention economy. The cross-layer configuration patterns are summarized in Table 3.

Table 3. Cross-Layer Structure of Strategic Archetype Families (Stimulus, Staging, and Response Synthesis)

Archetype Family	Core Stimulus Pattern (M1)	Core Staging Logic (M2)	Message Orientation (M3)	Dominant Response Climate (M4)	Discriminant Cross-Layer Feature
Immediate Musical/Aesthetic Loops	Close/medium proxemics; eye-level framing; direct gaze; tightly aligned audio–visual synchrony; high-arousal delivery	Arc 0 continuous alignment; payoff merges with hook; sustained expressive plateau	Emotional + entertainment values; parasocial immediacy; high-arousal	Rewatch- and save-oriented climates; affective return-behavior	Fusion of hook and payoff into a single instant multimodal plateau creating immediate recognizability
Proof-to-Reveal Configurations	Overlay-anchored premises; controlled framing; stepwise visual exposure; evidentiary cues	Arc 1 sequential progression; premise → evidence → reveal; punctual closure	Informational, demonstrative or credibility-oriented meanings; stance of showing	Save- and share-oriented climates; reference-driven circulation	Interpretive guidance through procedural clarification leading to evidentiary reveal
Process-to-CTA Funnels	Process-oriented visuals; multi-step craft or transformation cues; behind-the-scenes settings	Arc 1 or Arc 2 progression; reveal integrates CTA as narrative resolution	Transformational, instructional or world-building meanings; progression toward action	Conversion-oriented climates; intent-signaling comments; participation cues	CTA functions as the payoff, narrativizing action within the staging sequence
Escalation-to-Punchline Social Currency Formats	High-contrast or multi-setting cues; interpersonal tension; visually anomalous framings	Arc 2 escalation; complication → tension → recontextualizing punchline	Mixed-valence entertainment; anticipation, surprise or discomfort	Share-led climates; redistribution-driven circulation	Tension-Escalation Arcs resolved by a punchline that reorganizes prior cues

The first family, *Immediate Musical and Aesthetic Loops*, constructs a branding position of emotional accessibility and relational proximity. The artist appears embodied, present, and affectively open—not as a performer at a distance, but as someone the viewer is already close to. This position is produced through close- or medium-range proxemics, eye-level framing, direct gaze, cohesive settings, and tightly synchronized audio-visual timing. Gesture, beat, vocal onset, and gaze converge almost immediately, creating a continuous expressive plateau in which hook and payoff collapse into a single multimodal event. The underlying mechanism is experiential

synchrony: meaning is delivered through rhythmic and affective coherence, requiring little or no temporal elaboration. This configuration is well-suited to TikTok's interest graph logic, which prioritizes content that achieves interpretive clarity within the first few seconds. Save-led and rewatch-oriented response climates reflect the parasocial stance that they reinforce. Over time, emotional intensity, authenticity, and performative ease become recognizable elements of a persona—a stable branding grammar that viewers return to.

The second family, *Proof-to-Reveal Configurations*, positions the artist as credible, competent, and transparent. Visibility is secured through trustworthy demonstration—a branding stance that is structurally distinct from the affective pull of *Immediate Loops*. *Proof-to-Reveal* posts rely on controlled framing, premises anchored in overlays, and stepwise visual exposure that invite viewers into a procedural puzzle. Meaning is produced through an epistemic mechanism: interpretive uncertainty is reduced through sequential micro-evidence, culminating in a punctual reveal that reorganizes prior cues. This aligns well with TikTok's strong precedent for clarity-driven progression formats, which tend to reward posts that promise and deliver demonstrable results. Save-and-share-oriented response climates suggest that viewers come to treat these posts as reference points—the artist becomes someone whose process is worth consulting, not only experiencing.

The third family, *Process-to-CTA Funnels*, positions the artist as intentional, entrepreneurial, and agentic—someone with a trajectory the viewer is invited to join. These posts integrate process-oriented stimulus cues with transformational or instructional message orientations and action-centered narrative resolutions. The mechanism is distinct from that of Proof-to-Reveal: here, meaning is bound to the viewer's action. Sequential crafting, transformation, or world-building culminates in a payoff that serves as the call to action itself. Action becomes the logical resolution of the post's internal progression—a structural property that matters because CTA sequences embedded within narrative closure tend to be more effective than isolated prompts, given that the action is framed as a justified consequence of the preceding sequence. Conversion-oriented response climates reflect this: viewers interpret these posts as invitations into the artist's unfolding trajectory rather than as promotional interruptions.

The fourth family, *Escalation-to-Punchline Social Currency Formats*, positions the artist as memorable, socially intelligent, and capable of producing content that travels. For independent artists operating without established audiences, this is a structurally significant branding position: it is optimized for circulation across networks where the artist is not yet known. These posts begin

with interpretive ambiguity, develop rising tension, and resolve it via a punchline or recontextualization—with tension modulation as the core mechanism. Meaning is withheld until the payoff reorganizes the sequence, turning uncertainty into humor, surprise, or social insight. The viewer who feels the turn is the viewer who shares it. Share-driven circulation broadly supports this configuration, consistent with research suggesting that content eliciting surprise or incongruity tends to spread more readily, particularly when it functions as a socially portable token.

Across all four families, the archetypes address the competing demands of algorithmic legibility and relational credibility through distinct multimodal routes, each resolving the platform's visibility imperative through a different affective and narrative logic. Section 5.4 draws these findings together, connects them to the literature reviewed in Chapters 2 and 3, and develops the theoretical contribution they jointly make possible.

5.4 Discussion of Results

Across all three analytical layers, the findings converge on a single proposition: independent artists build visibility on TikTok by aligning stimulus-level cues, *micro-storytelling* mechanisms, and platform-visible engagement traces. That recurrence is not incidental—it reflects a collectively shared set of visibility-oriented practices, a grammar of legibility that cuts across genre, aesthetic register, and career stage (Bishop 2019, 2590–2596; Bonini & Gandini 2019, 9).

Interpreted through the *stimulus–staging–response* framework introduced in this study, and building on the S–O–R tradition (Mehrabian & Russell 1974), visibility emerges as a configuration-dependent property shaped by the interplay of perceptual cues, temporal organization, and audience engagement. Consistent with sociomaterial perspectives (Orlikowski & Scott 2008, 434–455), this process reflects the entanglement of platform infrastructures and creator practices—yet it remains empirically observable at the level of content, precisely through the alignment of stimulus, staging, and engagement-based response cues. The analytical separation of these layers serves a heuristic function: it enables precise examination of how perceptual cues derive meaning from their temporal positioning within sequences and how staging shapes the interpretation of engagement signals. This separation also exposes a structural gap that S–O–R models do not address: in platform contexts, response traces do not passively follow from stimuli but function as constitutive elements of the interpretive field through which content is evaluated and circulated—a reflexive quality that the *S–St–R* framework is designed to accommodate. As a configurational system, the framework captures

coordinated, relational interaction in short-form video contexts and specifies the conditions under which multimodal alignment produces platform legibility.

The RQ1 findings explain how the *paradox of independence*—shaped simultaneously by algorithmic gatekeeping, audience expectations of intimacy, and authenticity norms (see Chapter 2)—is negotiated at the level of content construction, a negotiation that becomes analytically legible through the recurrence of configural patterns across posts. Recognizability, then, is not a resource an independent artist accumulates over time and then draws on; it must be constructed within each post, requiring brand meaning to be re-established at the level of individual content instances—a demand that operates in tension with, and alongside, cumulative relationship development (Fournier 1998, 344). Each post must establish its own interpretive frame while contributing to broader identity coherence. Configuration logic provides the analytical mechanism through which this tension becomes tractable, specifying how episodic legibility and cumulative identity coherence operate as mutually reinforcing demands.

RQ2 extends this by introducing the temporal dimension: short-form storytelling constitutes a distinct communicative practice shaped by severe temporal constraints—one that shares structural features with longer narrative forms while operating within platform-specific constraints. *Micro-storytelling*, as discussed by Mueller and Rajaram (2022, 149), refers to short, platform-adapted forms of storytelling that fragment narratives across formats and sequences. These sequences resonate with foundational accounts of narrative structuring (Labov 1972, 362–370), which describe ordered phases such as orientation and resolution, but the present findings specify how such structuring is compressed under short-form conditions. Short-form video, as a context associated with rapid, low-effort, and affect-driven processing (Ma & Jiang 2024, 1–2), forecloses extended development, requiring key interpretive cues to be established early in the sequence; initial elements consequently function as primary orientation devices (Bouvier & Rasmussen 2022, 109–111). Stimulus cues acquire narrative force when organized into temporally ordered sequences, indicating that meaning construction depends on multimodal co-presence and, crucially, on the sequencing of those cues over time. This extends multimodal approaches to video analysis (Fazeli et al. 2023, 1–3) by foregrounding temporal organization as a critical condition for interpretability in short-form platform contexts.

Where RQ2 establishes the temporal and multimodal conditions governing individual sequences, RQ3 reveals that these conditions are not randomly distributed but organize into four higher-order configuration families. This finding advances configuration theory (Meyer et al. 1993, 1175–1176;

Fiss 2011, 393–394) by operationalizing configurational logic at the level of individual short-form content artifacts, extending its application beyond the organizational and strategy-focused contexts in which it has predominantly been developed. These archetypes differ in their narrative and affective pathways, yet each resolves the platform’s legibility imperative through a distinct multimodal configuration—demonstrating that legibility operates at the level of configurations rather than isolated features. The presence of hybrid and boundary cases confirms that these archetypes are recurring strategic tendencies, not prescriptive templates, and reflect the adaptive character of creator practice under algorithmically structured distribution conditions.

Response patterns, consistent with the SMEB literature (Dolan et al. 2016, 265; Muntinga et al. 2011, 15–16), differentiate these configurations while remaining analytically subordinate to meaning: they are observable traces of audience uptake, treated as contextual signals and not primary explanatory drivers, while still shaping the evaluative conditions under which content circulates. What this refines in the SMEB framework is the role of response mode: whereas Muntinga et al. (2011, 15–16) conceptualize engagement primarily in relation to user motivations and behavioral dispositions, the present findings demonstrate that response patterns are equally shaped by the multimodal configuration of the content itself—a relational dimension that variable-based engagement models have not theorised, and one that has direct consequences for how platform uptake is understood.

The findings constitute a theoretically grounded account of how visibility is constructed under metric-driven platform conditions—one most directly applicable to settings characterized by interest-graph distribution, high content velocity, and context collapse, where content must secure interpretability without the scaffold of prior audience familiarity. The recurrence of similar configurations across artists who differ markedly in genre, aesthetics, and audience scale is itself evidence of a shared configurational logic. By examining posts that have demonstrably achieved platform-visible uptake, the analysis isolates the multimodal conditions under which visibility is actively produced—and the consistency of those conditions across varied creative contexts confirms that content-level design is a structuring force in visibility outcomes, not an artefact of individual popularity or contextual circumstance.

The findings from RQ1, RQ2, and RQ3 converge on a unified set of conditions under which visibility is produced: RQ1 establishes the vocabulary of legibility at the level of stimulus cues; RQ2 demonstrates that this legibility is inherently temporal, requiring sequencing into expectation–resolution flows; and RQ3 reveals that these temporal dynamics crystallize into higher-order

configuration patterns. *Configuration fluency* names what these patterns share. Section 3.4 established it as a theoretical construct; the empirical findings specify its operational content: it is the capacity to assemble stimulus cues, temporal organization, and participatory affordances into configurations that are simultaneously coherent to viewers and legible to platform distribution logics. This fluency operates not as a stable trait but as a post-level accomplishment—a situated communicative competence that must be re-enacted with each content instance under algorithmically distributed, context-collapsed conditions. These three layers together form the study's enriched theoretical framework (Figure 4): an integrated account of how stimulus-level legibility, micro-temporal staging, and reflexive response dynamics jointly produce short-form visibility under interest-graph distribution. Chapter 6 develops its implications across engagement research, narrative theory, and sociomaterial accounts of platform visibility. In combination, these findings recast visibility as a configuration-dependent process rather than a variable-based outcome. Visibility is not something an actor possesses; it is continuously produced through multimodal alignment within individual content instances—actively constructed at the level of content design, not inherited from prior audience relationships or platform status.

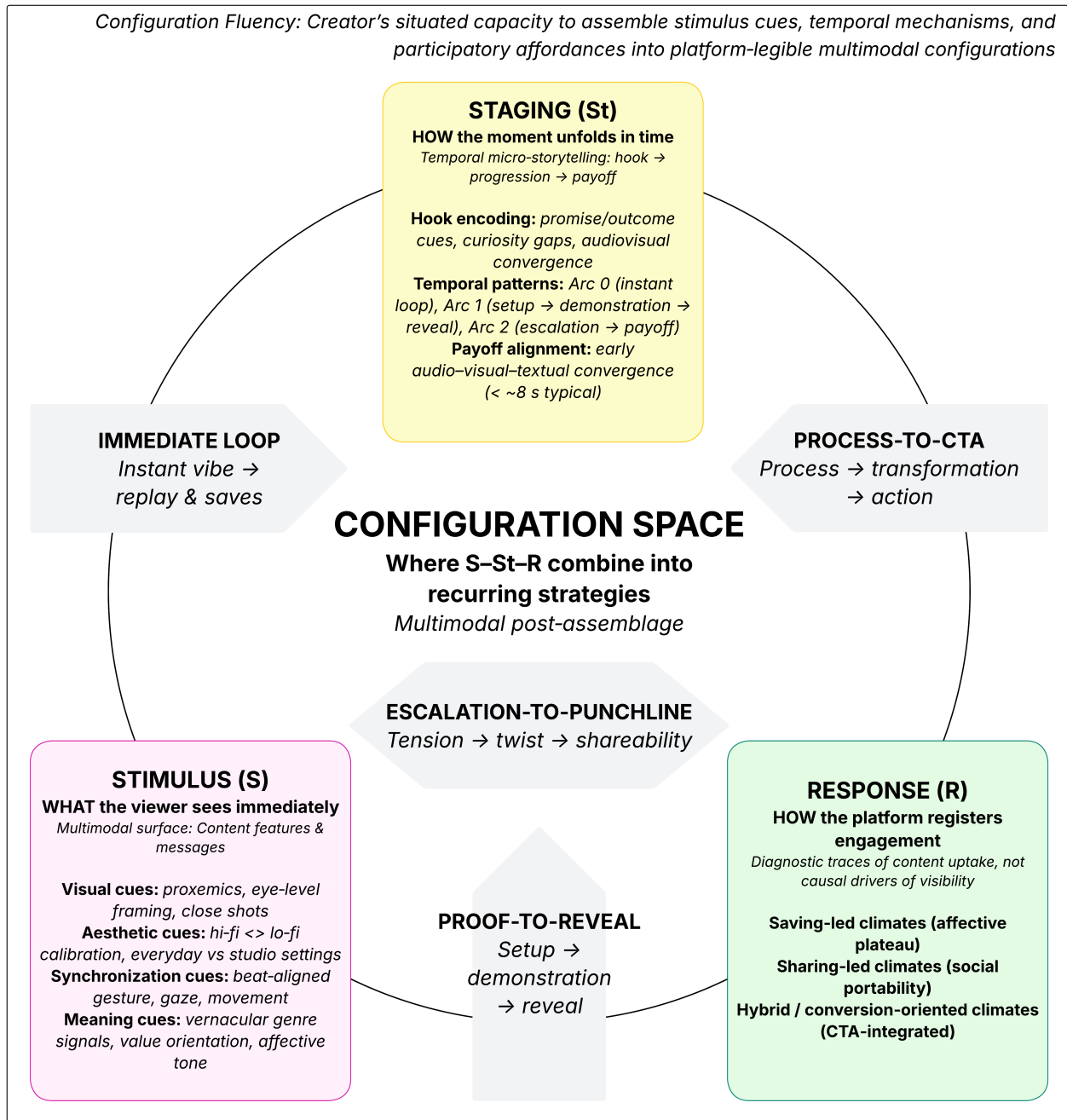


Figure 4. Configuration framework of stimulus, staging, and response in short-form video content

6 Conclusions

Chapter 5 established that independent artists construct visibility on TikTok through the systematic alignment of stimulus cues, temporal staging, and audience engagement—a finding whose theoretical, practical, and methodological implications this Chapter 6 develops. The recurrent multimodal configurations identified in the analysis extend, refine, and reposition existing theoretical understandings of artist branding, short-form narration, and platform-mediated visibility, reframing short-form posts as sociotechnical artifacts in which communicative meaning must be established locally within each post, under the structural pressures of context collapse, algorithmic opacity, and interest-graph distribution.

The chapter proceeds as follows. Section 6.1 develops the theoretical implications of the *configuration fluency* construct across three domains. Section 6.2 translates these insights into practical implications for independent artists. Section 6.3 reflects on the methodological contributions of integrating *Visual–Verbal Video Analysis* with the *stimulus–staging–response* logic. Section 6.4 outlines the study’s analytical boundaries and proposes directions for future research.

6.1 Theoretical Contributions

Short-form visibility on interest-graph platforms is a configuration-level phenomenon. The explanatory force lies not solely in individual content features, but in their combined and temporally coordinated effects across stimulus cues, temporal staging, and audience-oriented engagement signals—a shift that prior research, grounded in analytically separable variables and additive relationships between post features and engagement outcomes (Berger & Milkman 2012, 1–2; de Vries et al. 2012, 85–86; Xiao et al. 2026, 154, 168), has only begun to address. Three contributions follow—each addressing a domain where the present findings move the conversation forward.

The first contribution concerns what the right unit of explanation actually is. Drawing on configuration theory, which conceptualizes outcomes as emerging from internally coherent constellations of interdependent factors (Meyer et al. 1993, 1175–1176; Fiss 2011), recent studies have begun to acknowledge interaction effects between content and creator characteristics (Han et al. 2020, 576), yet dominant accounts remain grounded in analytically separable, variable-based approaches. The four *strategic archetypes* identified in Chapter 5 show directly how different yet internally coherent configurations can generate comparable high-engagement outcomes; in this sense, configuration—rather than individual features—is the unit of explanation.

The second contribution concerns not what is in a post, but what happens across it. As established in Section 3.3, prior work has developed the theoretical foundations of narrative staging—from classical narrative phases (Labov 1972) to platform-specific *micro-storytelling* forms (Mueller & Rajaram 2022, 25, 149–153)—yet stops short of specifying how narrative coherence is achieved under the extreme temporal compression of short-form video. This study proposes micro-temporal staging as a mechanism for structuring meaning and sustaining viewer engagement under these conditions. The three staging configurations—*Immediate Loops* (Arc 0), *Revealed Progressions* (Arc 1), and *Escalation Arcs* (Arc 2)—demonstrate how hooks, progression, and payoff are rearranged or compressed in short-form video. This account aligns with findings that emphasize rapid engagement, attention capture, and accelerated cognitive processing in short video environments (Ma & Jiang 2024, 1–2, 6–7), as well as research highlighting the ephemeral and loosely structured nature of social media discourse (Bouvier & Rasmussen 2022, 17).

The third contribution concerns where sociomaterial relations actually become visible—not at the level of platform infrastructure, but within individual posts. Building on prior work examining algorithmically informed gatekeeping (Bonini & Gandini 2019, 9) and creator practices under conditions of limited transparency (Bishop 2019, 2589–2590), the present study indicates that sociomaterial relations can be empirically traced within individual content items. In contexts where content circulates without prior audience–creator relationships, interpretability must be achieved locally through the coordination of multimodal elements, temporal sequencing, and interactional cues—combining sound, image, and gesture in platform-specific ways (Jung 2024, 243–244, 248). The recurring strategic patterns identified across cases suggest that these elements are consistently mobilized to support interpretability at the post level. In this way, sociomaterial entanglement is not only a background condition of platformed visibility but is enacted through the situated construction of individual content artifacts.

Configuration-level alignment, micro-temporal sequencing, and the legibility-first condition are analytically distinct yet mutually reinforcing mechanisms—each necessary for a full account of how content achieves visibility on interest-graph platforms. Together, the three contributions reorient the study of short-form visibility away from isolated variables and abstract platform structures toward the artifact-level processes through which visibility is actively produced. Next Section 6.2 translates these theoretical insights into practical implications for independent artists.

6.2 Managerial Contribution

A theory that cannot be acted on is incomplete. The contributions developed in Section 6.1 have direct practical bearing for independent artists navigating TikTok's interest-graph distribution infrastructure. *Configuration fluency* provides the central axis around which strategic content design can be oriented. Three implications follow, each grounded in the multimodal patterns observed across the high-engagement posts analyzed in Chapter 5.

The most consequential practical implication is a reorientation from expression-first to *legibility-first content design*. Unlike follower-based platforms, where audience familiarity reduces interpretive ambiguity, interest-graph distribution eliminates this buffer: posts appear before viewers with no relational history, and the video itself must provide all necessary cues for immediate comprehension. Tightly coordinated stimulus-level cues—close or medium framing, eye-level orientation, direct gaze, and immediate audio onset—jointly articulate the post's premise within the opening seconds, with save-led engagement climates reinforcing the importance of rapid clarity. A legibility-first approach ensures that expressive moves remain intelligible to unfamiliar audiences while preserving creative flexibility within a stable structural frame.

Timing is the second area of practical consequence. Coherence in short-form contexts emerges through micro-temporal sequencing, which requires practitioners to approach content creation as a series of temporal commitments: each moment is a structural decision, not merely an aesthetic one. In practical terms, emotional, musical, comedic, or demonstrative moments must occur at predictable junctures, with audio, gesture, expression, and textual cues working in concert. Calls to action are most effective when integrated into the payoff itself—as in the Process-to-CTA Funnel configuration—where the CTA becomes the natural resolution of an unfolding transformation, carrying the force of narrative consequence.

A third practical implication concerns the management of creative variation over time. Visibility emerges from the stabilization of patterned alignments, and the more durable strategic orientation is the consistency of the multimodal structure maintained across posts—not periodic reinvention. In saturated content conditions, creators often assume that visibility requires increasing production complexity—yet strategic distinctiveness emerges from variation within a consistent structural template. *Configuration fluency* reduces production burden by stabilizing the architectural backbone of posts while preserving room for surface-level creativity and artistic identity.

Collectively, these implications reposition strategic competence on TikTok: the practitioner's task is legible configuration work, not the decoding of opaque algorithmic mechanics. *Configuration fluency* provides the actionable foundation through which independent artists can reconcile the demands of visibility with limited resources—cultivating a consistent multimodal repertoire that meets the legibility conditions of interest-graph distribution while sustaining a credible artistic presence.

6.3 Methodological Contributions

Treating visuals, captions, sound, and engagement indicators as analytically separable domains—as much existing research does, isolating individual features for thematic interpretation or quantitative testing—forecloses the relational dynamics that make short-form content legible in the first place. This study suggests that short-form video content can be analyzed as a multimodal, temporally organized configuration in which meaning emerges from patterned co-presence across modalities. It indicates how these modalities can be integrated within a single analytical architecture that accounts for their interdependence under platform-native conditions—reorganizing and combining established multimodal tools into a framework attuned to TikTok's particular pressures of context collapse, temporal compression, and recommendation opacity.

The first methodological contribution arises from shifting the analytic focus from feature-based coding to configuration-based interpretation. Instead of treating cues such as shot scale, on-screen text, affective tone, or audio onset as separable predictors, the study positions them as interdependent components within broader *S–St–R* constellations. Integrating *Visual–Verbal Video Analysis* with the *stimulus–staging–response* logic developed in Chapter 3 enabled the independent coding of observable features, temporal sequencing decisions, and platform-visible engagement indicators, while still interpreting them as mutually shaping elements of the same multimodal configuration. This analytic shift proved essential in practice: the staging patterns that structure meaning in the dataset became visible only once the analytical architecture treated co-presence and timing as constitutive rather than incidental.

The second methodological contribution derives from the explicit integration of temporal sequencing into the analytical framework. Short-form videos are frequently coded for their content. The temporal unfolding of meaning has received comparatively little analytical attention—a gap this study addresses, and one the findings show to be more consequential than static content enumeration. Treating staging as its own analytical layer enabled segmentation into hooks, progressions, and payoffs, and the documentation of how visual, auditory, and textual cues realign

as each clip unfolds. Staging proves the most analytically generative of the three layers: it emerges not as embellishment applied to content already formed, but as a structural precondition for meaning-making in fast-scrolling feeds. The prevalence of *Immediate Loops* structures, the clustering of payoff moments, and the tight multimodal synchrony observed in hook windows all depended on treating temporal organization as constitutive of meaning rather than incidental to it.

The third methodological contribution concerns the specification of the unit of analysis. Treating the *post-assembly* as a bounded analytical unit maintains coherence across video content, paratextual elements, and visible engagement traces without collapsing these components into a single category. Audiovisual cues, on-screen text, captions, and platform-visible responses are treated as co-present elements that jointly shape meaning within a flat, mutually constitutive structure. Comments, shares, and saves were incorporated as contextual signals characterizing response climates, with normalized ratios and a $\geq 1.2\times$ dominance threshold distinguishing among save-led, share-led, and comment-led patterns. This threshold was deliberately conservative—its purpose is to keep a clear boundary between observable engagement outputs and unobservable user motivations or underlying platform mechanisms, a distinction that social media research has frequently collapsed.

The fourth methodological contribution concerns the analytical generalization procedure through which observed patterns were elevated to configuration-level archetypes. By requiring each archetype to recur across multiple artists and in more than one post per artist, the study grounds its higher-order categories in demonstrably patterned phenomena. This recurrence criterion is one of the framework's most consequential design decisions. It is what distinguishes the archetypes reported in Chapter 5 from post-hoc interpretive constructs, anchoring them instead in stable, cross-case sociotechnical regularities.

Across these four contributions, the study suggests that qualitative analysis of short-form video can account for multimodality, temporality, and platform mediation without relying on opaque proprietary datasets or reducing content to decontextualized feature lists. The framework—*VVA* integrated with *S-St-R* logic and anchored in the *post-assembly* unit—offers a transparent methodological pathway for tracing evolving platform vernaculars, shifting engagement patterns, and emerging multimodal conventions. Its replicability across platforms, time periods, and content categories is a methodological contribution in itself. Whether the specific *configuration logics* identified here hold beyond TikTok's particular conditions is among the most productive questions for future work.

6.4 Limitations and Future Research

The scope of this study is defined by its analytical commitment to publicly observable *post-assemblages* and by the epistemologically grounded decision to center interpretation on platform-visible data. This boundary is both practical and principled: by anchoring analysis within the content artifact itself, the study maintains a clear epistemological distinction between publicly accessible evidence and the private analytics, internal recommendation-system information, and production processes that constitute separate objects of inquiry. The study advances a bounded empirical perspective on how meaning, legibility, and engagement become visible on the multimodal surface of short-form posts, treating TikTok's algorithmic infrastructure as a contextual condition rather than an object of direct investigation. That precision is the boundary's purpose, not its cost. Future research could extend this perspective by combining artifact-level interpretation with ethically obtained exposure or retention metrics—such as watch-time curves or completion data—to examine how temporal structuring interacts with attention dynamics while maintaining a clear distinction between observable patterns and unobservable platform mechanisms.

A second analytical boundary concerns the purposive focus on high-engagement posts created by independent artists. This sampling strategy was designed to enable the identification of stable *S–St–R* configuration patterns in contexts where visibility had demonstrably been achieved, providing a theoretically grounded foundation for comparative work. The deliberate focus on high-engagement content means that low- and mid-engagement posts, as well as content produced under different resource conditions or within different industry infrastructures, constitute productive terrain for subsequent investigation. Future studies could investigate whether the archetypes identified here generalize across varying levels of visibility, or whether posts that do not achieve traction fragment, recombine, and shift in ways that do. This comparison is among the most productive avenues this work opens. Larger-scale analyses using automated multimodal extraction or machine-learning-assisted segmentation could examine the prevalence and distribution of *S–St–R* configurations. At the same time, comparative work across genres, creator roles, national platform contexts, and gendered positions could clarify how platform vernaculars and resource conditions shape configuration practices. Cross-platform analyses could further test whether *configuration logics* remain stable across platforms with differing feed architectures, cultural norms, or paratextual conventions.

A third analytical consideration concerns the single-coder design. Systematic procedures were applied throughout to support analytic consistency—including a structured coding framework,

recoding checks, and a maintained audit trail—and the interpretation of multimodal, temporally unfolding content is explicitly grounded in researcher judgment as a constitutive feature of qualitative inquiry. The complexity of analyzing stimulus-level cues, temporal staging sequences, and multimodal alignment positions, cross-coder validation is a productive methodological extension rather than a correction. Future research could strengthen the robustness of the analytical framework by incorporating multi-coder adjudication, testing *VVA*-based extraction matrices with independent analysts, or developing hybrid workflows in which computational tools pre-segment audiovisual events before human interpretation. Such work could clarify which elements of the *S–St–R* framework exhibit stable cross-coder reliability and which components rely more heavily on contextual expertise, abductive reasoning, or theoretical saturation.

Finally, the study’s focus on artist branding as externally observable configuration work reflects a deliberate analytical commitment to the *post-assemblage* as the primary unit of inquiry. By centering analysis at this level, the study achieves precision in accounting for the multimodal surface of content artifacts. Creators’ stage production processes, longitudinal artistic trajectories, and cross-platform promotional strategies constitute generative directions for subsequent inquiry. Tracing how configuration patterns emerge, stabilize, or erode over time at the level of the individual artist represents the most generative work. Future studies could integrate artifact-level analysis with interviews, platform-native ethnography, or longitudinal creator tracking to examine how artists adapt, maintain, or abandon specific configurations in response to evolving incentive structures. Such research could also illuminate how *configuration fluency* develops across a career and how short-form practices intersect with broader processes of audience development, identity construction, and cultural participation.

The analytical boundaries outlined here collectively define a research agenda, not a set of constraints. The framework’s grounding in publicly observable *post-assemblages* makes it replicable and extensible—across platforms, time periods, and content categories—precisely because it does not depend on proprietary data or platform access. The configuration-oriented logic introduced here offers a conceptual basis for examining how visibility, meaning, and participation are organized within platform-native communication—questions whose significance is likely to grow as interest-graph distribution becomes an increasingly dominant condition of cultural participation. Stepping back from the framework level to the level of everyday artistic practice makes the stakes of these questions more concrete. The 99,000 daily tracks that opened this thesis keep coming—and the artists behind them are still asking the same question: how do you get heard? This study does not resolve that question, but it reframes it. Visibility is not something that happens

to a post; it is something a post makes happen. Understanding how is where the work of helping independent artists begins.

References

- Abidin, Crystal (2016) “Aren’t these just young, rich women doing vain things online?”: Influencer selfies as subversive frivolity. *Social Media + Society*, Vol. 2 (2).
- Abidin, Crystal (2017) #familygoals: Family influencers, calibrated amateurism, and justifying young digital labor. *Social Media + Society*, Vol. 3 (2).
- Adaval, Rashmi – Wyer, Robert S. (1998) The role of narratives in consumer information processing. *Journal of Consumer Psychology*, Vol. 7 (3), 207–245.
- Akdevelioglu, Duygu – Venkatesh, Alladi (2022) Consumer ties and social media consumer culture in social media networks. In: *The SAGE handbook of social media marketing*, eds. Annmarie Hanlon – Tracy Tuten, 58–71. SAGE Publications Ltd, London.
- Alexander, Julia (2019) Old Town Road proves TikTok is a new SoundCloud for artists. *The Verge* 5.4.2019. <https://www.theverge.com/2019/4/5/18296815/lil-nas-x-old-town-road-tiktok-artists-spotify-soundcloud-streams-revenue>, accessed 31 March 2026.
- ALLEA (2023) *The European Code of Conduct for Research Integrity – Revised Edition 2023*. Berlin.
- Ambady, Nalini – Rosenthal, Robert (1992) Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, Vol. 111 (2), 256–274.
- Ameh, Samuel S. – Emmanuel, N. O. (2025) The TikTok effect: How storytelling in short-form videos shape discourses on refugees in Canada – A mixed methods approach. *Howard Journal of Communications*, 1–15.
- Appel, Gil – Grewal, Lauren – Hadi, Rhonda – Stephen, Andrew T. (2020) The future of social media in marketing. *Journal of the Academy of Marketing Science*, Vol. 48 (1), 79–95.
- Arriagada, Arturo – Bishop, Sophie (2021) Between commerciality and authenticity: The imaginary of social media influencers in the platform economy. *Communication, Culture and Critique*, Vol. 14 (4), 568–586.
- Ashley, Christy – Tuten, Tracy (2015) Creative strategies in social media marketing: An exploratory study of branded social content and consumer engagement. *Psychology & Marketing*, Vol. 32 (1), 15–27.
- Batra, Rajeev – Ahtola, Olli T. (1990) Measuring the hedonic and utilitarian sources of consumer attitudes. *Marketing Letters*, Vol. 2 (2), 159–170.
- Baym, Nancy K. (2018) *Playing to the crowd: Musicians, audiences, and the intimate work of connection*. NYU Press, New York.

- Benham, Tom (2026) Music metadata made easy: Why it matters and how to get it right. *Identity Music*. <https://identitymusic.com/blog/music-metadata-made-easy>, accessed 31 March 2026.
- Berger, Jonah – Milkman, Katherine L. (2012) What makes online content viral? *Journal of Marketing Research*, Vol. 49 (2), 192–205.
- Bhandari, A. – Bimo, S. – Noone, R. – Butkowski, C. P. (2022) What does the internet sound like? Panel presented at AoIR 2022: The 23rd Annual Conference of the Association of Internet Researchers. <http://spir.aoir.org>, accessed 31 March 2026.
- Bishop, Jonathan (2005) Building international empires of sound: Concentrations of power and property in the “global” music market. *Popular Music and Society*, Vol. 28 (4), 443–471.
- Bishop, Sophie (2019) Managing visibility on YouTube through algorithmic gossip. *New Media & Society*, Vol. 21 (11–12), 2589–2606.
- Boddy, Clive R. (2016) Sample size for qualitative research. *Qualitative Market Research: An International Journal*, Vol. 19 (4), 426–432.
- Bonini, Tiziano – Gandini, Alessandro (2019) “First week is editorial, second week is algorithmic”: Platform gatekeepers and the platformization of music curation. *Social Media + Society*, Vol. 5 (4), 1–11.
- Bösch, Miriam – Divon, Tamar (2024) The sound of disinformation: TikTok, computational propaganda, and the invasion of Ukraine. *New Media & Society*, Vol. 26 (9), 5081–5106.
- Bouncken, Ricarda B. – Czakon, Wojciech – Schmitt, Florian (2026) Purposeful sampling and saturation in qualitative research methodologies: Recommendations and review. *Review of Managerial Science*, Vol. 20 (2), 579–615.
- Bouvier, Gwen – Rasmussen, Joel (2022) *Qualitative research using social media*. Routledge, London.
- boyd, danah (2011) Social network sites as networked publics: Affordances, dynamics, and implications. In: *A networked self*, ed. Zizi Papacharissi, 47–66. New York: Routledge.
- Breitman, Kira (2024) TikTok video size guide: Best dimensions for 2025. Riverside 12.12.2024. <https://riverside.com/blog/tiktok-video-size>, accessed 31 March 2026.
- Burr, Vivien (1995) *An introduction to social constructionism*. Routledge, London.
- Caswell, Estelle – Daniels, Michael (2022) We tracked what happens after TikTok songs go viral. YouTube. <https://www.youtube.com/watch?v=S1m-KgEpoow>, accessed 31 March 2026.
- Ceci, Laura (2025) TikTok – statistics & facts. Statista. <https://www.statista.com/topics/6077/tiktok/#topicOverview>, accessed 31 March 2026.

- Chahal, Hardeep – Wirtz, Jochen – Verma, Anshu (2019) Social media brand engagement: Dimensions, drivers and consequences. *Journal of Consumer Marketing*, Vol. 37 (2), 191–204.
- Chen, K.-J. – Lin, J.-S. – Choi, J. H. – Hahm, J. M. (2015) Would you be my friend? An examination of global marketers' brand personification strategies in social media. *Journal of Interactive Advertising*, Vol. 15 (2), 97–110.
- 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*, Vol. 23 (2), e21463.
- Cheng, Z. – Li, Y. (2024) Like, comment, and share on TikTok: Exploring the effect of sentiment and second-person view on user engagement with TikTok news videos. *Social Science Computer Review*, Vol. 42 (1), 201–223.
- Cheung, M. F. Y. – To, W. M. (2016) Service co-creation in social media: An extension of the theory of planned behavior. *Computers in Human Behavior*, Vol. 65, 260–266.
- Chow, Andrew R. (2019) Inside Lil Nas X's record-breaking, culture-changing summer. *TIME*. <https://time.com/5652803/lil-nas-x-2/>, accessed 31 March 2026.
- Content Marketing Institute (n.d.) What is content marketing? <https://contentmarketinginstitute.com/what-is-content-marketing/>, accessed 5.2.2025.
- Copyright Service (n.d.) Copyright and phonographic copyright. https://copyrightservice.co.uk/copyright/copyright_symbol, accessed 30.3.2026.
- Corbin, Juliet – Strauss, Anselm (2008) *Basics of qualitative research: Techniques and procedures for developing grounded theory*. 3rd ed. SAGE Publications, Los Angeles.
- Cotter, Kelley (2019) Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. *New Media & Society*, Vol. 21 (4), 895–913.
- Curmi, Francesco (2022) Social contagion and virality in online networks. In: *The SAGE handbook of social media marketing*, eds. Annmarie Hanlon – Tracy Tuten, 146–160. SAGE Publications Ltd, London.
- Danilova, K. – Krupa, O. (2021) Brand-management theoretical principles in music business. *Socio-Cultural Management Journal*, Vol. 4 (2), 50-78.
- David, M. E. – Roberts, James A. (2025) TikTok brain: An investigation of short-form video use, self-control, and phubbing. *Social Science Computer Review*, Vol. 43 (6), 1292–1305.
- de Vries, L. – Gensler, S. – LeeFlang, Peter S. H. (2012) Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of Interactive Marketing*, Vol. 26 (2), 83–91.

- Divon, Tamar – Ebbrecht-Hartmann, Tobias (2025) “On TikTok, everything needs to be playful, even the Holocaust!”: Playability, memes, and participatory memory culture. *New Media & Society*, 1–32.
- Dolan, Rebecca – Conduit, Jodie – Fahy, John – Goodman, Steve (2016) Social media engagement behavior: A uses and gratifications perspective. *Journal of Strategic Marketing*, Vol. 24 (3–4), 261–277.
- Dolan, Rebecca – Conduit, Jodie – Frethey-Bentham, Catherine – Fahy, John – Goodman, Steve (2019) Social media engagement behavior. *European Journal of Marketing*, Vol. 53 (10), 2213–2243.
- Dzedzickis, Andrius – Kaklauskas, Artūras – Bucinskas, Vytautas (2020) Human emotion recognition: Review of sensors and methods. *Sensors*, Vol. 20 (3), 592.
- Eadon, Y. M. (2024) “You could hear a hair pin drop”: Queer utopianism and informal knowledge production in the Gaylor closeting conspiracy theory. *Social Media + Society*, Vol. 10 (2).
- Edlom, Jessica (2022) The engagement imperative: Experiences of communication practitioners’ brand work in the music industry. *Media and Communication*, Vol. 10 (1), 66–76.
- Eriksson, Päivi – Kovalainen, Anne (2008) *Qualitative methods in business research*. SAGE Publications Ltd, London.
- Escalas, Jennifer E. – Stern, Barbara B. (2003) Sympathy and empathy: Emotional responses to advertising dramas. *Journal of Consumer Research*, Vol. 29 (4), 566–578.
- Fazeli, S. – Sabetti, J. – Ferrari, Mario (2023) Performing qualitative content analysis of video data in social sciences and medicine: The visual-verbal video analysis method. *International Journal of Qualitative Methods*, Vol. 22, 1–17.
- Fiss, Peer C. (2011) Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, Vol. 54 (2), 393–420.
- Fournier, Susan (1998) Consumers and their brands: Developing relationship theory in consumer research. *Journal of Consumer Research*, Vol. 24 (4), 343–353.
- Frenneaux, Richard (2023) The rise of independent artists and the paradox of democratisation in the digital age: Challenges faced by music artists in the new music industry. *DIY, Alternative Cultures & Society*, Vol. 1 (2), 125–137.
- Gibbs, Martin – Meese, James – Arnold, Michael – Nansen, Bjorn – Carter, Marcus (2015) #funeral and Instagram: Death, social media, and platform vernacular. *Information, Communication & Society*, Vol. 18 (3), 255–268.

- Gligor, David – Bozkurt, S. – Russo, Ivan (2019) Achieving customer engagement with social media: A qualitative comparative analysis approach. *Journal of Business Research*, Vol. 101, 59–69.
- Green, Melanie C. – Brock, Timothy C. (2000) The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, Vol. 79 (5), 701–721.
- Grzenkiewicz, Martyna – Wildfeuer, Janina (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*, Vol. 40 (4), 1143–1166.
- Guest, Greg – Namey, Emily – Chen, Mitchell (2020) A simple method to assess and report thematic saturation in qualitative research. *PLOS ONE*, Vol. 15 (5), e0232076.
- Han, Y. – Lappas, Theodoros – Sabnis, Gaurav (2020) The importance of interactions between content characteristics and creator characteristics for studying virality in social media. *Information Systems Research*, Vol. 31 (2), 576–588.
- Hanlon, Annmarie – Tuten, Tracy L. (2022) Introduction to social media marketing. In: *The SAGE handbook of social media marketing*, eds. Annmarie Hanlon – Tracy L. Tuten, 3–13. London: SAGE Publications.
- Hansen, Lars K. – Arvidsson, Adam – Nielsen, Finn Å. – Colleoni, Elanor – Etter, Michael (2011) Good friends, bad news – Affect and virality in Twitter. In: *Future information technology, Pt II*, eds. L. T. Yang – J. J. Park – C. Lee, 34–43. Berlin Heidelberg: Springer.
- Haynes, Jo – Marshall, Lee (2018) Beats and tweets: Social media in the careers of independent musicians. *New Media & Society*, Vol. 20 (5), 1973–1993.
- He, A.-Z. – Cai, Y. – Cai, L. – Zhang, Y. (2021) Conversation, storytelling, or consumer interaction and participation? The impact of brand-owned social media content marketing on consumers' brand perceptions and attitudes. *Journal of Research in Interactive Marketing*, Vol. 15 (3), 419–440.
- Hennink, Monique – Kaiser, Bonnie N. (2022) Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, Vol. 292, 114523.
- Hirschman, Elizabeth C. – Holbrook, Morris B. (1982) Hedonic consumption: Emerging concepts, methods and propositions. *Journal of Marketing*, Vol. 46 (3), 92–101.
- Hodgson, Thomas (2021) Spotify and the democratisation of music. *Popular Music*, Vol. 40 (1), 1–17.
- Hoffmann, Anna Lauren – Jonas, Andreas (2016) Recasting justice for internet and online industry research ethics. In: *Internet research ethics for the social age: New cases and challenges*, eds. Michael Zimmer – Katharina Kinder-Kurlanda, 3–18. New York: Peter Lang.

- Horton, Donald – Wohl, Richard (1956) Mass communication and para-social interaction. *Psychiatry*, Vol. 19 (3), 215–229.
- Hracs, Brian J. (2012) A creative industry in transition: The rise of digitally driven independent music production. *Growth and Change*, Vol. 43 (3), 442–461.
- Hutchinson, Andrew (2022) TikTok confirms that 10 minute video uploads are coming to all users. *Social Media Today*. <https://www.socialmediatoday.com/news/tiktok-confirms-that-10-minute-video-uploads-are-coming-to-all-users/619535/>, accessed 31 March 2026.
- Ingeird, Helene – Fossheim, Hallvard (2016) Introductory remarks. In: *Internet research ethics*, eds. Helene Ingeird – Hallvard Fossheim, 9–13. Oslo: Cappelen Damm Forskning.
- Jensen Schau, Hope – Gilly, Mary C. (2003) We are what we post? Self-presentation in personal web space. *Journal of Consumer Research*, Vol. 30 (3), 385–404.
- Johar, J. S. – Sirgy, M. Joseph (1991) Value-expressive versus utilitarian advertising appeals: When and why to use which appeal. *Journal of Advertising*, Vol. 20 (3), 23–33.
- Jung, Bernhard (2024) Traveling sounds, embodied responses: Aesthetic reflections on TikTok. In: *Traveling music videos*, eds. Tomáš Jirsa – Mathias Bonde Korsgaard, 243–260. Bloomsbury Academic, London.
- Kakkuri-Knuuttila, Marja-Liisa (2006) *Mitä on tutkimus? Argumentaatio ja tieteenfilosofia*. Gaudeamus, Helsinki.
- Kress, Gunther – van Leeuwen, Theo (2020) *Reading images*. Routledge, London.
- Labov, William (1972) The transformation of experience in narrative syntax. In: *Language in the inner city: Studies in the Black English vernacular*, 354–396. University of Pennsylvania Press, Philadelphia.
- Lacher, Katherine T. (1989) Hedonic consumption: Music as a product. In: *Advances in Consumer Research*, Vol. 16, 367–373. Association for Consumer Research.
- Lacher, Katherine T. – Mizerski, Richard (1994) An exploratory study of the responses and relationships involved in the evaluation of, and in the intention to purchase new rock music. *Journal of Consumer Research*, Vol. 21 (2), 366–372.
- Lapid, Lyn (2021) “I heard this sound makes the photos in your camera roll aesthetic 🤩!?! #fyp #inmymind #ThatCloseMessenger”. TikTok video 21.9.2021. <https://www.tiktok.com/@lynlapid/video/7010188269325438214>, accessed 31 March 2026.
- Ledin, Per – Machin, David (2018) Film clips. In: *Doing visual analysis: From theory to practice*, 133–162. SAGE Publications Ltd, London.

- Lewis, Geoffrey J. – Graham, Gordon – Hardaker, Glenn (2005) Evaluating the impact of the internet on barriers to entry in the music industry. *Supply Chain Management: An International Journal*, Vol. 10 (5), 349–356.
- Li, B. – Chong, Alex Y.-L. – Ch, E. – Yee-Loong, A. (2015) What triggers sharing in viral marketing? The role of emotion and social feature. In: *Proceedings of PACIS 2015*. <http://aisel.aisnet.org/pacis2015/95>, accessed 31 March 2026.
- Li, D. – Li, W. – Lu, B. – Li, H. – Ma, S. – Krishnan, G. – Wang, J. (2025) Delving deep into engagement prediction of short videos. In: *Computer vision – ECCV 2024*, Vol. 15112, eds. A. Leonardis – E. Ricci – S. Roth – O. Russakovsky – T. Sattler – G. Varol, 289–306. Springer Nature Switzerland.
- Li, Y. – Guan, M. – Hammond, Philip – Berrey, Lauren 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*, Vol. 36 (3), 261–271.
- Ling, Cheng – Blackburn, Jeremy – de Cristofaro, Emiliano – Stringhini, Gianluca (2022) Slapping cats, bopping heads, and Oreo shakes: Understanding indicators of virality in TikTok short videos. In: *Proceedings of the 14th ACM Web Science Conference 2022*, 164–173.
- Lipschultz, Jeremy H. (2020) Social crowds, voice, and personal branding. In: *Social media measurement and management: Entrepreneurial digital analytics*, 3–34. Routledge, New York.
- Lu, Y. – Shen, Cindy (2023) Unpacking multimodal fact-checking: Features and engagement of fact-checking videos on Chinese TikTok (Douyin). *Social Media + Society*, Vol. 9 (1).
- Luminate (2024) *Luminate year-end report*. Luminate.
- Ma, Lin – Jiang, Q. (2024) Swiping more, thinking less: Using TikTok hinders analytic thinking. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, Vol. 18 (3).
- Malterud, Kirsti – Siersma, Volkert D. – Guassora, Ann Dorrit (2016) Sample size in qualitative interview studies. *Qualitative Health Research*, Vol. 26 (13), 1753–1760.
- Malthouse, Edward C. – Haenlein, Michael – Skiera, Bernd – Wege, E. – Zhang, Michael (2013) Managing customer relationships in the social media era: Introducing the social CRM house. *Journal of Interactive Marketing*, Vol. 27 (4), 270–280.
- Martins, Gustavo B. – Papa, João P. – Adeli, Hojjat (2020) Deep learning techniques for recommender systems based on collaborative filtering. *Expert Systems*, Vol. 37 (6).
- Marwick, Alice E. (2013) *Status update: Celebrity, publicity, and branding in the social media age*. Yale University Press, New Haven.

- Mayring, Philipp (2014) *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. <https://www.beltz.de>, accessed 31 March 2026.
- Mehrabian, Albert – Russell, James (1974) Conceptual framework. In: *An approach to environmental psychology*, 4–8.
- Meyer, Alan D. – Tsui, Anne S. – Hinings, C. R. (1993) Configurational approaches to organizational analysis. *Academy of Management Journal*, Vol. 36 (6), 1175–1195.
- Mills, A. J. – John, J. (2025) Brand stories: Bringing narrative theory to brand management. *Journal of Strategic Marketing*, Vol. 33 (1), 35–53.
- Mitev, Nathalie – Howcroft, Debra (2011) Post-structuralism, social shaping of technology, and actor-network theory: What can they bring to IS research? In: *The Oxford handbook of management information systems: Critical perspectives and new directions*, eds. Robert Galliers – Wendy Currie, 292–322. Oxford: Oxford University Press.
- Mohammed, I. – Mahmoud, M. A. – Hinson, Robert E. (2025) Short video narratives and Porter's framework: The role of parasocial interaction. *Tourism Critiques: Practice and Theory*, Vol. 6 (2), 205–224.
- Morgan, Robert M. – Hunt, Shelby (1999) Relationship-based competitive advantage. *Journal of Business Research*, Vol. 46 (3), 281–290.
- Morris, Jeremy W. (2020) Music platforms and the optimization of culture. *Social Media + Society*, Vol. 6 (3).
- Mueller, Milton E. – Rajaram, D. (2022) Understand and implement mobile story genres. In: *Social media storytelling*, 147–219. London: Routledge.
- Muntinga, Daniël G. – Moorman, Marjolein – Smit, Edith G. (2011) Introducing COBRAs. *International Journal of Advertising*, Vol. 30 (1), 13–46.
- Murphy, Nora A. – Hall, Jeffrey A. (2021) Capturing behavior in small doses: A review of comparative research in evaluating thin slices for behavioral measurement. *Frontiers in Psychology*, Vol. 12.
- Murphy, S. – Hume, Margee (2025) The new digital music marketing ecosystem: Artist direct. *Creative Industries Journal*, Vol. 18 (2), 172–204.
- Mustikawati, R. – Sadewa, G. P. – Fadholi, M. A. (2023) Vertical video trends among amateur digital platform users as an alternative for film production. *Journal of Urban Society's Arts*, Vol. 9 (2), 129–136.
- Nastisin, L. – Mruz, O. – Gavura, S. – Mruzova, Y. N. (2025) A lexical and sentiment analysis of TikTok hooks and their influence on digital engagement. *Marketing and Management of Innovations*, Vol. 16 (4), 64.

- Nelson-Field, Karen – Riebe, Erica – Newstead, Karyn (2013) The emotions that drive viral video. *Australasian Marketing Journal*, Vol. 21 (4), 205–211.
- Nickell, Christopher (2020) Promises and pitfalls: The two-faced nature of streaming and social media platforms for Beirut-based independent musicians. *Popular Communication*, Vol. 18 (1), 48–64.
- Nikolinakou, Angeliki – King, Karen W. (2018) Viral video ads: Emotional triggers and social media virality. *Psychology & Marketing*, Vol. 35 (10), 715–726.
- Orlikowski, Wanda J. – Scott, Susan V. (2008) Sociomateriality: Challenging the separation of technology, work and organization. *Academy of Management Annals*, Vol. 2 (1), 433–474.
- Packard, Mark D. (2017) Where did interpretivism go in the theory of entrepreneurship? *Journal of Business Venturing*, Vol. 32 (5), 536–549.
- Palombo, Nicholas G. (2024) “Sexual context collapse” on TikTok: Platform politics, content moderation, and user agency in platform governance. Master’s thesis, York University. <https://yorkspace.library.yorku.ca/items/a40bc454-cb6a-42d0-92eb-b94b37b8f926>, accessed 31 March 2026.
- Phillips, Barbara J. – McQuarrie, Edward F. – Griffin, William G. (2014) How visual brand identity shapes consumer response. *Psychology & Marketing*, Vol. 31 (3), 225–236.
- Pilipets, Elena (2021) Hashtagging, duetting, sound-linking: TikTok gestures and methods of (in)distinction. *The Journal of Media Art Study and Theory*, Vol. 4, 109–135.
- Pivecka, N. – Ratzinger, R. A. – Florack, Arnd (2022) Emotions and virality: Social transmission of political messages on Twitter. *Frontiers in Psychology*, Vol. 13, 1–8.
- Polletta, Francesca – Chen, P. C. B. – Gardner, B. G. (2011) The sociology of storytelling. *Annual Review of Sociology*, Vol. 37, 109–130.
- Portugal, I. – Alencar, P. – Cowan, D. (2018) The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems with Applications*, Vol. 97, 205–227.
- Pulla, Vasilios – Carter, E. (2018) Employing interpretivism in social work research. *International Journal of Social Work and Human Services Practice*, Vol. 6 (1), 9–14.
- Radovanović, Bojana (2022) TikTok and sound: Changing the ways of creating, promoting, distributing and listening to music. *INSAM Journal of Contemporary Music, Art and Technology*, No. 9, 51–73.
- Radovanović, V. (2024) Independent music marketing: Trends and insights 2024. Videobolt Blog 21.6.2024. <https://blog.videobolt.net/post/independent-music-marketing-trends-insights-research-report-2024>, accessed 31 March 2026.

- Rodas-Coloma, A. – Cabezas-González, M. – Casillas-Martín, S. – Moreno, P. N.-B. (2026) Quality vs. populism in short video political communication: A multimodal study of TikTok. *Journalism and Media*, Vol. 7 (1), 46.
- Rogers, Richard (2021) After vanity metrics: Critical analytics for social media analysis. In: *Research methods for digital work and organization: Investigating distributed, multi-modal, and mobile work*, 271–290. Oxford University Press, Oxford.
- Rowley, Jennifer (2008) Understanding digital content marketing. *Journal of Marketing Management*, Vol. 24 (5–6), 517–540.
- Russell, James A. – Clement, J. – Jiwani, D. – Ridgeway, K. – Schroeder, M. (1980) A circumplex model of affect. *Journal of Personality and Social Psychology*, Vol. 39 (6), 1161–1178.
- Ryan, Gemma (2018) Introduction to positivism, interpretivism and critical theory. *Nurse Researcher*, Vol. 25 (4), 14–20.
- Salles, J. (2025) Affect and prediction in short video social media recommendation algorithm: TikTok and the missing half-second. *New Media & Society*.
- Sashi, C. M. (2012) Customer engagement, buyer-seller relationships, and social media. *Management Decision*, Vol. 50 (2), 253–272.
- Savic, Milica (2021) From Musical.ly to TikTok: Social construction of 2020's most downloaded short video app. *International Journal of Communication*, Vol. 15, 3173–3194.
- Schank, Roger C. (1999) *Dynamic memory revisited*. Cambridge University Press, Cambridge.
- Scherer, Klaus R. (2005) What are emotions? And how can they be measured? *Social Science Information*, Vol. 44 (4), 695–729.
- Schreier, Margrit (2024) *Qualitative content analysis in practice*. SAGE Publications Ltd, London.
- Schreiner, M. – Fischer, T. – Riedl, René (2021) Impact of content characteristics and emotion on behavioral engagement in social media: Literature review and research agenda. *Electronic Commerce Research*, Vol. 21 (2), 329–345.
- Serazio, Michael (2024) *The authenticity industries: Keeping it “real” in media, culture, and politics*. Stanford: Stanford University Press.
- Shay, R. – van der Horst, M. (2019) Using brand equity to model ROI for social media marketing. *International Journal on Media Management*, Vol. 21 (1), 24–44.
- Sinha, N. – Ahuja, V. – Medury, Y. (2011) Corporate blogs and internet marketing – Using consumer knowledge and emotion as strategic variables to develop consumer engagement. *Journal of Database Marketing & Customer Strategy Management*, Vol. 18 (3), 185–199.

- Sood, Suruchi – Rogers, Everett M. (2000) Dimensions of parasocial interaction by letter-writers to a popular entertainment-education soap opera in India. *Journal of Broadcasting & Electronic Media*, Vol. 44 (3), 386–414.
- Spotify for Artists (2025) Glossary of music industry terms.
<https://artists.spotify.com/blog/glossary-of-music-industry-terms>, accessed 31 March 2026.
- Stassen, Murray (2024) Only 19% of artists on Spotify had over 1,000 monthly listeners in 2023. *Music Business Worldwide*. <https://www.musicbusinessworldwide.com/only-19-of-artists-on-spotify-had-over-1000-monthly-listeners-in-2023/>, accessed 31 March 2026.
- Stewart, S. M. (2018) The development of the artist-fan engagement model.
<https://musicbusinessresearch.wordpress.com>, accessed 31 March 2026.
- Suominen, K.-A. (2023) Artificial intelligence in research.
- Tadesse, Tsion (2025) Global Recorded Music Revenues Reach \$36.2 Billion in 2024 Amid Slower Growth. MIDia Research. <https://www.midiaresearch.com/blog/global-recorded-music-revenues-reach-362-billion-in-2024-amid-slower-growth>, accessed 1 April 2026.
- Taecharungroj, V. – Nueangjamnong, P. (2015) Humor 2.0: Styles and types of humor and virality of memes on Facebook. *Journal of Creative Communications*, Vol. 10 (3), 288–302.
- Tafesse, Wondwesen – Wood, Benjamin P. (2021) Followers’ engagement with Instagram influencers: The role of influencers’ content and engagement strategy. *Journal of Retailing and Consumer Services*, Vol. 58, 102303.
- Tao, Y. – Sun, S. – He, J. (2026) Unraveling the multimodal features and audience responses to thinspiration videos: A computational analysis of Chinese TikTok (Douyin). *Body Image*, Vol. 56, 102053.
- TikTok (n.d.-a) Camera tools. <https://support.tiktok.com/en/using-tiktok/creating-videos/camera-tools>, accessed 12 December 2025.
- TikTok (n.d.-b) Duets. <https://support.tiktok.com/en/using-tiktok/creating-videos/duets>, accessed 13 March 2025.
- TikTok (n.d.-c) Stitch. <https://support.tiktok.com/en/using-tiktok/creating-videos/stitch>, accessed 13 March 2025.
- Timmermans, Stefan – Tavory, Iddo (2012) Theory construction in qualitative research. *Sociological Theory*, Vol. 30 (3), 167–186.
- Tuten, Tracy L. (2024) *Social media marketing*. London: SAGE.
- University of Turku (n.d.) Artificial intelligence in teaching and studying.
- van Dijck, José (2009) Users like you? Theorizing agency in user-generated content. *Media, Culture & Society*, Vol. 31 (1), 41–58.

- Vaughn, Richard (1980) How advertising works: A planning model. *Journal of Advertising Research*, Vol. 20 (5), 27–33.
- Vinerean, Simona (2017) Content marketing strategy: Definition, objectives and tactics. *Expert Journal of Marketing*, Vol. 5 (2), 92–98.
- Vizcaino-Verdu, A. – Abidin, Crystal (2022) Music challenge memes on TikTok: Understanding in-group storytelling videos. *International Journal of Communication*, Vol. 16, 883–908.
- Wang, Z. – Kim, H. G. (2017) Can social media marketing improve customer relationship capabilities and firm performance? Dynamic capability perspective. *Journal of Interactive Marketing*, Vol. 39, 15–26.
- Winner, Langdon (1993) Upon opening the black box and finding it empty: Social constructivism and the philosophy of technology. *Science, Technology, & Human Values*, Vol. 18 (3), 362–378.
- Xiao, L. – Li, X. – Mou, J. (2026) Exploring user engagement behavior with short-form video advertising on short-form video platforms: A visual-audio perspective. *Internet Research*, Vol. 36 (1), 154–188.
- Xiao, L. – Sun, L. – Lei, Z. – Fu, B. (2026) How do features in health short-form videos impact viewer engagement: An empirical study. *Industrial Management & Data Systems*, Vol. 126 (2), 678–709.
- 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*, Vol. 20 (2), 69.
- Yang, Y. – Yu, B. – Yu, G. (2024) Cascade or updown sliding: The influence of interactive interface of short video app on information processing and user experience. *International Communication of Chinese Culture*, Vol. 11 (2), 115–132.
- Zhang, Z. – Qiu, K. – Ye, Y. (2025) Influence of audiovisual features of short video advertising on consumer engagement behaviors: Evidence from TikTok. *Journal of Business Research*, Vol. 201, 115662.
- Zhang, Z. – Zhang, L. (2025) Most significant impact on consumer engagement: An analytical framework for the multimodal content of short video advertisements. *Journal of Theoretical and Applied Electronic Commerce Research*, Vol. 20 (2), 54.
- Zhao, H. – Wagner, Christian (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*, Vol. 33 (2), 820–849.

- Zhou, R. (2024) Understanding the impact of TikTok's recommendation algorithm on user engagement. *International Journal of Computer Science and Information Technology*, Vol. 3 (2), 201–208.
- Zhu, C. – Hall, C. Michael – Fong, L. H. N. – She, S. – Naderi Koupaei, S. (2024) How does ice-snow celebrity influence Chinese tourists to visit Nordic destinations? The case of TikTok. *Scandinavian Journal of Hospitality and Tourism*, Vol. 24 (1), 106–122.
- Zhu, C. – Xu, X. – Zhang, W. – Chen, J. – Evans, Richard (2019) How health communication via TikTok makes a difference: A content analysis of TikTok accounts run by Chinese provincial health committees. *International Journal of Environmental Research and Public Health*, Vol. 17 (1), 192.
- Zimmer, Michael (2010) “But the data is already public”: On the ethics of research in Facebook. *Ethics and Information Technology*, Vol. 12 (4), 313–325.
- Zulli, David – Zulli, David J. (2022) Extending the internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media & Society*, Vol. 24 (8), 1872–1890.
- Zummo, Maria L. (2022) Young generation and accessibility to health dissemination. *Journal of Language and Discrimination*, Vol. 6 (1), 93–113.

Appendices

Appendix 1 VVVA Analytical Protocol

This appendix documents the technical Visual–Verbal Video Analysis (VVVA) procedure and provides the operational and procedural details that underpin the analytical framework presented in Sections 4.5–4.6. It consolidates the complete VVVA workflow for extracting multimodal evidence from the dataset.

The *VVVA* method involves a rigorous six-step analytical framework: (1) collecting, organizing, and reviewing data; (2) transcribing verbal data; (3) choosing units of analysis; (4) extracting and coding data; (5) organizing, describing, and interpreting extracted data; and (6) reporting findings (Fazeli et al. 2023, 4–5).

Segmentation and boundaries used:

- The hook window is defined as 0–3 seconds, with allowance for sub-second pre-roll. A hook must establish the initial expectation visually, verbally, or sonically.
- Progression is coded by identifying the highest coherent sequencing structure visible in the post (e.g., Setup→Payoff or Setup→Complication→Payoff).
- The payoff is timestamped at the moment the post delivers its principal narrative, emotional, musical, or informational closure.
- Audio–visual relationships are coded as Aligned, Complementary, or Divergent based on sustained evidence across the post.

Multimodal transcription conventions used:

- Speech and lyrics are transcribed verbatim where legible.
- Gestural cues such as gaze, posture, and salient micromovements are annotated with approx. ± 0.5 s precision.
- Onscreen text is captured with exact wording and placement.
- Audio source (TikTok library/original sound) is logged, including whether the platform explicitly indicates artist ownership of the audio.

Appendix 2 Chartmetric Discovery Procedure

This appendix documents the full technical details of the discovery protocol underpinning the sample construction described in Section 4.3.

Discovery procedure: The dataset was identified between 16–17 June 2025 using Chartmetric’s public Artist List. The search targeted creators with demonstrated TikTok activity, emphasizing two-month engagement rate, career stage (developing/mid-level/mainstream), and label type (large indies/other/self-released). These filters were selected to maximize information power while retaining independence diversity. Creators with inactive or sporadic posting behavior during 2025 were excluded, as such profiles risked distorting engagement comparisons.

Independence criterion: Independence was operationalized via Spotify’s public ℙ/© release metadata, as detailed in Section 4.3. Cases with missing, conflicting, or ambiguous ownership markers were excluded.

Documentation: All inclusion and exclusion decisions are documented in the controlled-access case index (Appendix 3).

Appendix 3 Anonymized Case Index & Metadata

The creators in the dataset are anonymized as Female Artist 1–5 and Male Artist 1–5. Metadata includes the capture date, post duration, video format, and post age (in days) at the time of capture. The full case index is available to examiners on request via the controlled-access link detailed in the Data Access Controls section below.

Anonymization approach: Anonymization procedures follow the principles detailed in Sections 4.3 and 4.8. Because TikTok orders comments by ranking rather than chronology, visible comments are treated only as bounded contextual traces, not as representative samples.

Data access controls: A secure, read-only mapping linking anonymized IDs to original TikTok URLs is stored in University OneDrive and is available only to examiners on request; further redistribution is prohibited.

Appendix 4 VVVA Coding Manual (Matrices 1–4)

This appendix provides the complete operational specification of the coding instrument—including variables, allowed values, decision rules, and boundary criteria. Sections 4.5 and 4.6 address the methodological rationale and abductive development of these operationalizations. See Appendix 7 for row-level instantiations of these decision rules.

How to read: Each matrix contains a single compact table listing variables, allowed values, and the key decision rule. The table is the auditable snapshot of the instrument.

Table 4. Coding Manual Matrix 1 for Content Features and Visual Form

Variable	Allowed Values	Decision Rule/Exceptions
Duration	integer	Record visible length
Format	Standard/Duet/Stitch/Comment-reply	UI must indicate format; else 'unclear'.
Production value	Raw/Lo-Fi; Polished/Hi-Fi	Default Raw unless lighting+steadiness+audio jointly justify Hi-Fi.
Setting count	1/2/3+	Distinct environments by visual cues.
Main setting	Home/Studio/Street/Venue/Other	Dominant environment anchoring activity.
Dominant POV	1st/2nd/3rd/Mixed/Unclear	Dominance sampling: 1-s intervals; $\geq 50\%$ = dominant; 40–50% = Mixed.
Shot scale	Close/Medium/Long/Mixed/Unclear	Dominance sampling as above.
Camera angle	High/Eye/Low/Mixed/Unclear	Dominance sampling as above.
On-screen text	None/Overlay/Subtitles/Both	Environmental signage excluded.
Caption function	Question/Context/Decorative	Assign the function most lost if caption removed.
Hashtag type	Generic/Niche/Mixed	Based on audience-narrowing effect.
Mentions (dominant)	Venue/Collaborator/Brand/Fellow Artist/Media/Other	By presence weight.
Audio source	TikTok library/Original sound/None	As labeled at capture.
Audio ownership	Artist's own (confirmed)/Other's (confirmed)/Unclear	Only if explicitly indicated.

Table 5. Coding Manual Matrix 2a for Narrative Flow and Audio

Variable	Allowed Values	Decision Rule/Exceptions
Hook (0–3 s) description	Free text	Functional leeway if brief pre-roll; must set initial expectation.
Hook modality	Audio-led/Visual-led/Aligned	Aligned = both channels indispensable for same event.
Hook type (dominant)	Trend/Context; Curiosity/Question; Surprise/Pattern break; Promise/Outcome	Choose dominant expectation logic.
Secondary hook logic	Same list or None	Use only if co-dominant and structurally persistent.
Progression arc	0=Loop; 1=Setup→Payoff; 2=Setup→Complication→Payoff; 3=≥4 phases	Prefer the highest structure that yields expectation-release.
Payoff type (dominant)	Musical; Reveal/Twist; Punchline; Lesson/Tips; Emotional beat; CTA payoff	Timestamp first dominant onset; annotate sustained fulfilment if applicable.
Secondary payoff element	Same list or None	Use if clearly co-dominant.
Payoff modality	Audio-led/Visual-led/Aligned	As per hook modality.
CTA present/location	Yes/No; Overlay/Speech/Both	Minor end-tag CTA ≠ payoff unless conversion is endpoint.
Audio–visual relationship	Aligned/Complementary/Divergent	Whole-post judgement based on dominant relation.

Table 6. Coding Manual Matrix 2b for Progression Arc Configurations

Arc Code	One-line description	Typical timestamping
0 Loop	Chorus loop performance; no structural change	No discrete payoff timestamp
1 Setup→Payoff	Teaser text 'wait for it...' → drop/reveal	Payoff ~0:07–0:12
2 Setup→Complication→Payoff	Initial cue → minor obstacle/twist → resolved hit	Payoff ~0:10–0:15
3 ≥4 phases	Orientation → escalation → turn → closure	First payoff onset marked; sustained fulfilment noted

Table 7. Coding Manual Matrix 3 for Content Messages, Characters, and Affect

Variable	Allowed Values	Decision Rule/Exceptions
Topic (primary)	Performance/preview; Behind-the-scenes; Relatability; Relationship/social; Tutorial/tips; Trend participation; Aesthetic/lore; Promotion/announcement; Community interaction; Other	Dominant aboutness after full viewing.
Vernacular genre	Storytelling/Confessional; Skit/Comedy; Performance; Tutorial/Demo; Vlog/Diary; Trend/Template; Other	Platform-native delivery convention; format ≠ genre.
Value type (primary)	Entertainment; Informational; Emotional; Transformational; Transactional	Value realised at payoff; assign single dominant.
Value type (secondary)	Same list or None	Use only if persistent hybridity.
Payload driver (dominant)	Musical payoff; Humor; Personal story; Relatability; CTA	Mechanism carrying the value.
Affect: Valence	Positive/Negative/Neutral/Mixed	Dominant at payoff window; spikes don't reclassify.
Affect: Arousal	High/Medium/Low	Sustained dominance.
Primary character type	Human/Animal/Object	—
Artist as primary	Yes/No	—
Secondary characters	None/Present (count)	Mentions without representation do not count.
Vulnerability cue	Yes/No	Mark Yes only with an observable self-exposure move.
Reciprocity prompt	Yes/No	Explicit invitation to participate.

Table 8. Coding Manual Matrix 4 for Engagement Response and Strategic Synthesis

Variable	Allowed Values	Decision Rule/Exceptions
Raw counts	Views, Likes, Comments, Shares, Saves	Snapshot at capture; retain capture date.
Per-view ratios	comments/views; shares/views; saves/views	Descriptive normalisation for comparability.
High-effort ERV	(comments + shares + saves)/views	Descriptive intensity cue; not a benchmark.
Engagement mode	Comment-led/Share-led/Save-led/Balanced	Rule: if top ratio $\geq 1.2\times$ the next, assign that mode; otherwise Balanced.
Reception climate	Supportive/Contested/Adversarial/Spammy	Bounded by ranking effects; paraphrase themes.
Visibility/Conversion/Relational cue intensity	Low/Medium/High	Count-based descriptive indices.
Audio strategy	Discovery-led/Conversion-led/Hybrid	Assign with minimum two converging cues.
Archetype label & memo	Free text	Assign only with cross-matrix warrant and cross-artist recurrence.

Extended Worked Examples

Example 1: Trend-template adaptation

- Stimulus: Medium close-up with on-screen text establishing a curiosity-based opening. TikTok library audio is used.
- Staging: Hook (0–2 s) via question overlay; rapid cuts escalate into a reveal/twist payoff around 0:09; audio-visual alignment is sustained.
- Message: Topic = trend participation; value = entertainment (primary); dominant affect = positive, medium arousal.
- Response pattern: Typically share-led, as portable punchlines travel well across audiences.

Example 2: Tutorial micro-demo

- Stimulus: Over-the-shoulder POV of DAW screen with subtitles; original sound with voiceover.

- Staging: Setup→Payoff with a compact lesson/tip delivered around 0:11; audio-visual relation is complementary.
- Message: Topic = tutorial; value = informational; affect = neutral–positive, low–medium arousal.
- Response pattern: Often save-led, as instructional content is stored for later use.

Decision and Boundary Rules:

- Single-location rule: Each observation is recorded only once in its primary matrix; cross-references prevent double-counting. (Rationale: Section 4.4 and 4.6.)
- Ambiguity resolution: Functional priority—what sets the expectation and what fulfills it.
- Affect coding: Determined at the payoff window; transients do not override sustained affect.
- Engagement mode: The dominant per-view ratio must be $\geq 1.2 \times$ the next-highest to qualify; otherwise coded as Balanced. (Rationale: Section 4.5.)
- Audio-ownership: Logged only when the interface explicitly confirms it.

Appendix 5 Data Management Plan

- **Description of Research Data:** The research data for this thesis consists of secondary, publicly available observational data. It includes: (1) a discovery dataset containing public Chartmetric metadata used for case selection; (2) an anonymized case index linking internal IDs to public TikTok video URLs; and (3) the primary analytical dataset, comprising detailed Excel coding matrices (codebooks, multimodal transcriptions, and quantitative engagement snapshots) generated by the researcher during the analysis phase.
- **Permissions and Data Protection:** Because the study relies exclusively on publicly accessible social media artifacts and platform-surfaced metadata, no formal research permits or interview consent forms were required. Although the data is public, the dataset includes identifiers (creator handles and video URLs). To comply with ethical guidelines and mitigate privacy risks, all creator identities, specific video links, and verbatim top comments have been anonymized or paraphrased in the final thesis report.
- **Storage, Security, and Version Control:** Throughout the research process, all data files (including the raw case index and the working Excel matrices) were stored securely on a password-protected University of Turku (UTU) provided cloud storage service (UTU OneDrive), accessible only to the researcher and the supervisory team. Version control was strictly maintained by keeping the raw, unmodified data index separate from the active, working analysis matrices. The coding matrices inherently served as a research diary, documenting the stages of analysis, coding decisions, and analytical memos.
- **Post-Research Retention and Destruction:** The researcher assumes full responsibility for the data after the completion of the thesis. To allow for academic verification and the standard grading appeal period, the complete, raw analytical dataset (Excel matrices and case index) will be securely retained in UTU-approved storage for 6 months following the official grading and approval of the Master's thesis. After this retention period expires, all associated digital files, raw data, and URLs will be permanently destroyed.

Appendix 6 Explanation of the Use of AI

- **Scope of AI Usage:** In accordance with the University of Turku (UTU) guidelines (Suominen, 2023; University of Turku, n.d.), this section details the application of generative AI and AI-assisted tools during the research and writing process of this thesis. AI tools were used solely as support tools for data processing, language refinement, and conceptual structuring. AI was not used to generate the core analytical findings, nor is it cited as a primary academic source or author.
- **Data Processing and Transcription:** AI-assisted transcription software was used to generate initial text drafts of the audio and spoken elements within the TikTok video corpus. To strictly adhere to UTU data protection guidelines and ensure research integrity, the data-sharing and model-training settings within these tools were explicitly disabled before any audio files were processed, preventing the proprietary research dataset from being absorbed into external language models.
- **Information Retrieval and Ideation:** Generative AI language models (e.g., ChatGPT, Copilot) and AI-integrated database search functions (e.g., RAG search) were used during the preliminary phases of the research to brainstorm search terms, structure complex theoretical concepts, and explore broader contexts. However, all factual claims, theoretical frameworks, and previous literature incorporated into the thesis were independently verified, sourced, and cited using traditional, peer-reviewed academic databases.
- **Language Editing and Proofreading:** Generative AI tools were employed to translate non-English data excerpts and to edit language throughout the writing process, including grammar checking, improving sentence flow, reducing redundancy, and ensuring a consistent academic tone. All AI-assisted suggestions were critically evaluated and manually revised by the researcher before inclusion in the final manuscript.
- **Academic Integrity and Responsibility:** As mandated by the UTU ethical guidelines for learning and research, the researcher assumes full and final responsibility for the content, correctness, analysis, and referencing of this thesis.

Appendix 7 Row-Level Synthesis Tables (RQ1–RQ2 Audit Trail)

This appendix consolidates the per-post evidence from the January–June 2025 corpus referenced in Chapter 5. The tables list, for each anonymized post, the coded hook, arc, and payoff variables (Matrices 1–3), payoff timing (Matrix 2), and engagement-mode calculations (Matrix 4). Engagement modes are assigned using the $\geq 1.2\times$ dominance rule specified in Appendix 4: a post is classified as save-led, share-led, or comment-led when the dominant per-view ratio exceeds the next highest ratio by at least 20%; otherwise, the post is coded as balanced. Assignment is based on normalized per-view ratios captured at the time of data collection and is therefore interpretive rather than absolute. Variable definitions, boundary criteria, and procedural decision rules are specified in Appendix 4, while numeric source values are drawn directly from the coding workbook.

Table 9. Configuration Synthesis and Engagement Modes by Post

Post ID	Arc (M2)	Hook Type (M2)	Payoff Type (M2)	Payoff Modality (M2)	Configuration Family	Value (M3)	Affect (M3)	Engagement Mode (M4)	Archetype Label
Female artist 1__1	0	Surprise/Pattern-break	Musical	Aligned	Aesthetic Immersion Loop	Entertainment	Mixed, High	Save-led	Viral Dance Catalyst
Female artist 1__2	1	Promise–Outcome	Reveal/Twist	Aligned	Competence Demonstration	Entertainment	Mixed	Save-led	Misheard-Lyric Proof Hook
Female artist 1__3	0	Surprise/Pattern-break	Musical	Aligned	Aesthetic Immersion Loop	Entertainment	High	Share-led	Chemistry Showcase
Female Artist 2__1	0	Curiosity/Question	Musical	Aligned	Aesthetic Immersion Loop	Emotional	High	Save-led	Unreleased Tease Magnet
Female Artist 2__2	0	Curiosity/Question	Musical	Aligned	Aesthetic Immersion Loop	Emotional	Neutral	Save-led	Comparison Controversy Node
Female Artist 2__3	0	Curiosity/Question	Musical	Aligned	Aesthetic Immersion Loop	Emotional	High	Save-led	Countdown Converter
Female Artist 3__1	0	Promise–Outcome	Reveal/Twist	Visual-led	Aesthetic Immersion Loop	Emotional	Positive, High	Save-led	Aesthetic Relatability (Hair POV)
Female Artist 3__2	1	Promise–Outcome	Reveal/Twist + CTA	Visual-led	Integrated Commercial Funnel	Transformational	Positive, High	Save-led	Lore-to-Product Funnel
Female Artist 3__3	0	Surprise/Pattern-break	Emotional beat	Aligned	Aesthetic Immersion Loop	Emotional	Positive, High	Save-led	Mythic Muse Aesthetic
Female artist 4__1	0	Curiosity/Question	Musical	Aligned	Aesthetic Immersion Loop	Transformational	Positive, High	Save-led	Discovery-to-Fandom Bridge
Female artist 4__2	2	Curiosity/Question	Punchline	Aligned	Relational Social Currency	Entertainment/Informational	Negative, High	Save-led	Conflict/Street Narrative

Female artist 4__3	1	Curiosity/ Question	Reveal/ Twist + CTA	Visual-led	Integrated Commercial Funnel	Emotional	Mixed, High	Save-led	Guerrilla Marketing Builder
Female Artist 5__1	1	Promise– Outcome	Musical	Audio-led	Competence Demonstration	Informational	Neutral	Share-led	Tech Educator Performer
Female Artist 5__2	1	Promise– Outcome	Musical	Audio-led	Competence Demonstration	Informational	Neutral	Save-led	Algorithmic Music Curiosity
Female Artist 5__3	1	Promise– Outcome	Lesson/ Tips + Musical	Aligned	Competence Demonstration	Informational	Neutral, High	Save-led	Niche-to-Mainstre am Translator
Male Artist 1__1	0	Promise– Outcome	Musical	Audio-led	Aesthetic Immersion Loop	Emotional	Positive, High	Save-led	Emotional Comfort Artist
Male Artist 1__2	2	Promise– Outcome	Reveal/ Twist + CTA	Aligned	Competence Demonstration	Informational	Neutral	Save-led	BTS Bond Builder
Male Artist 1__3	1	Curiosity/ Question	Reveal/ Twist + Musical	Aligned	Competence Demonstration	Emotional	Mixed	Save-led	Origin Story Keeper
Male Artist 2__1	0	Promise– Outcome	Musical	Audio-led	Aesthetic Immersion Loop	Transformatio nal	Mixed, High	Save-led	Faith-Genre Fulfiller
Male Artist 2__2	0	Promise– Outcome	Musical	Audio-led	Aesthetic Immersion Loop	Transformatio nal	Positive	Save-led	Devotional Partner Ideal
Male Artist 2__3	0	Promise– Outcome	Musical	Audio-led	Aesthetic Immersion Loop	Entertainment	Positive, High	Save-led	FYP Discovery Darling
Male Artist 3__1	1	Surprise/Patter n-break	Punchline	Divergent	Relational Social Currency	Entertainment	Mixed	Share-led	Heartthrob Rediscovery
Male Artist 3__2	1	Promise– Outcome	Punchline	Aligned	Relational Social Currency	Entertainment	Neutral	Share-led	Fan-Driven Collab Engine
Male Artist 3__3	0	Trend/Context signal	Punchline	Aligned	Relational Social Currency	Entertainment	Mixed	Share-led	Chaotic Creative Persona

Male Artist 4__1	1	Promise– Outcome	Musical + Reaction	Aligned→ Compl.	Competence Demonstration	Entertainment	Mixed	Save-led	Studio Reaction Performer
Male Artist 4__2	1	Question + Surprise	Punchline	Aligned	Relational Social Currency	Entertainment	Negative	Save-led	Debate Bait Loop
Male Artist 4__3	1	Promise– Outcome + Surprise	Punchline	Aligned	Relational Social Currency	Entertainment	Mixed	Save-led	Meme-Fueled Persona
Male Artist 5__1	1	Promise– Outcome	Musical	Aligned	Competence Demonstration	Entertainment	Mixed	Share-led	Cover Transformer
Male Artist 5__2	1	Promise– Outcome	Musical	Aligned	Competence Demonstration	Entertainment	Mixed	Save-led	Fandom Meme Distributor
Male Artist 5__3	1	Curiosity/ Question	Musical (late)	Audio-led	Competence Demonstration	Entertainment	Mixed	Save-led	Grief-to-Strength Vocalist

Table 10. Content Messages and Staging Characteristics

Post ID	Topic (M3)	Vernacular Genre (M3)	Value	Affect	Hook Modality (M2)	Arc (M2)	Payoff Type	Payoff Modality	CTA Present
Female artist 1__1	Promotion/ announcement	Performance	Entertainment	Mixed–High	Aligned	0	Musical	Aligned	No
Female artist 1__2	Trend participation	Performance	Entertainment	Mixed	Visual-led	1	Reveal	Aligned	No
Female artist 1__3	Trend participation	Performance	Entertainment	High	Aligned	0	Musical	Aligned	No
Female Artist 2__1	Community interaction	Storytelling/ Confessional	Emotional	High	Aligned	0	Musical	Aligned	No
Female Artist 2__2	Music performance/p review	Performance	Emotional	Neutral	Aligned	0	Musical	Aligned	No
Female Artist 2__3	Promotion/ announcement	Performance	Emotional	High	Aligned	0	Musical	Aligned→ Compl.	No
Female Artist 3__1	Tutorial/tips	Tutorial/Demo	Emotional	Positive–High	Visual-led	0	Reveal	Visual-led	No
Female Artist 3__2	Aesthetic/ lore + Promotion	Vlog/Diary	Transformational	Positive–High	Visual-led	1	Reveal + CTA	Visual-led	Yes
Female Artist 3__3	Behind-the-scenes/ process	Vlog/Diary	Emotional	Positive–High	Aligned	0	Emotional beat	Aligned	No
Female artist 4__1	Music performance/p review	Performance	Transformational	Positive–High	Aligned	0	Musical	Aligned	No
Female artist 4__2	Trend participation	Storytelling/ Confessional	Entertainment/ Informational	Negative–High	Aligned	2	Punchline	Aligned	No
Female artist 4__3	Promotion/ announcement	Vlog/Diary	Emotional	Mixed–High	Visual-led	1	Reveal + CTA	Visual-led	Yes

Female Artist 5__1	Tutorial/tips	Tutorial/Demo	Informational	Neutral	Aligned	1	Musical	Audio-led	No
Female Artist 5__2	Tutorial/tips	Tutorial/Demo	Informational	Neutral	Aligned	1	Musical	Audio-led	No
Female Artist 5__3	Tutorial/tips	Tutorial/Demo	Informational	Neutral–High	Aligned	1	Lesson/Ti ps + Musical	Aligned	No
Male Artist 1__1	Promotion/ announcement	Vlog/Diary	Emotional	Positive–High	Audio-led	0	Musical	Audio-led	No
Male Artist 1__2	Music performance/p review	Performance	Informational	Neutral	Aligned	2	Reveal + CTA	Aligned	Yes
Male Artist 1__3	Relatability/ everyday life	Vlog/Diary	Emotional	Mixed	Aligned	1	Reveal + Musical	Aligned	No
Male Artist 2__1	Promotion/ann ouncement	Vlog/Diary	Transformational	Mixed–High	Audio-led	0	Musical	Audio-led	No
Male Artist 2__2	Music performance/p review	Performance	Transformational	Positive	Audio-led	0	Musical	Audio-led	No
Male Artist 2__3	Community interaction	Vlog/Diary	Entertainment	Positive–High	Aligned	0	Musical	Aligned	No
Male Artist 3__1	Relatability/ everyday life	Vlog/Diary	Entertainment	Mixed	Aligned	1	Punchline	Divergent	No
Male Artist 3__2	Relatability/ everyday life	Vlog/Diary	Entertainment	Neutral	Aligned	1	Punchline	Aligned	No
Male Artist 3__3	Relationship/s ocial dynamics	Vlog/Diary	Entertainment	Mixed	Trend/Context	0	Punchline	Aligned	No
Male Artist 4__1	Promotion/ announcement	Performance	Entertainment	Mixed	Aligned	1	Musical + Reaction	Aligned→ Compl.	No
Male Artist 4__2	Trend participation	Vlog/Diary	Entertainment	Negative	Aligned	1	Punchline	Aligned	No

Male Artist 4__3	Relatability/ everyday life	Vlog/Diary	Entertainment	Mixed	Aligned	1	Punchline	Aligned	No
Male Artist 5__1	Trend participation	Vlog/Diary	Entertainment	Mixed	Aligned	1	Musical	Aligned	No
Male Artist 5__2	Community interaction	Performance	Entertainment	Mixed	Aligned	1	Musical	Aligned	No
Male Artist 5__3	Relatability/ everyday life	Vlog/Diary	Entertainment	Mixed	Audio-led	1	Musical (late)	Audio-led	No

Table 11. Payoff Timing and Resolution Cues

Post ID	Arc	Payoff Timestamp (s)	Window	Resolution Cue
Female artist 1__1	0	~2	Early	Aligned musical onset
Female artist 1__2	1	~4	Early	Aligned reveal
Female artist 1__3	0	~8	Early	Aligned dance drop
Female Artist 2__1	0	~9	Early	Aligned vocal line
Female Artist 2__2	0	~15	Early	Aligned vocal line
Female Artist 2__3	0	~24	Early→Sustain	Aligned vocal peak
Female Artist 3__1	0	~25	Early→Sustain	Visual reveal
Female Artist 3__2	1	~38	Mid→End	Visual reveal + CTA
Female Artist 3__3	0	~50	Late sustain	Aesthetic beat
Female artist 4__1	0	~76	Early→Sustain	Aligned lyric onset
Female artist 4__2	2	~91	End	Punchline
Female artist 4__3	1	~38	End	CTA integrated
Female Artist 5__1	1	~end	End	Audio drop
Female Artist 5__2	1	~end	End	Audio drop
Female Artist 5__3	1	~end	End	Filter demo
Male Artist 1__1	0	~2	Early	Audio onset
Male Artist 1__2	2	~55	End	Verbal CTA
Male Artist 1__3	1	~15	Mid	Reveal + music
Male Artist 2__1	0	~2–6	Early	Aligned lyric onset
Male Artist 2__2	0	~2–6	Early	Aligned lyric onset

Male Artist 2__3	0	~0-2	Early	Aligned lyric onset
Male Artist 3__1	1	~end	End	Visual punchline
Male Artist 3__2	1	~end	End	Facial punchline
Male Artist 3__3	0	~early	Early	Rapid comedic beats
Male Artist 4__1	1	~early→mid	Early/Mid	Musical + reaction
Male Artist 4__2	1	~2	Early	Punchline
Male Artist 4__3	1	~2	Early	Punchline
Male Artist 5__1	1	~4	Early	Audio transform
Male Artist 5__2	1	~4	Early	Audio transform
Male Artist 5__3	1	~50	End	Late instrumental drop

Table 12. Engagement Metrics and Dominant Modes for Matrix 4

Post ID	Views	Comments	Shares	Saves	Comments/Views	Shares/Views	Saves/Views	Mode
Female artist 1__1	35,880,000	16,803	101,548	147,600	0.000468	0.00283	0.004114	Save-led
Female artist 1__2	27,610,000	7,791	32,446	177,400	0.000282	0.001175	0.006425	Save-led
Female artist 1__3	18,000,000	7,688	30,300	7,960	0.000427	0.001683	0.000442	Share-led
Female Artist 2__1	325,000	748	1,520	15,400	0.0023	0.004675	0.047361	Save-led
Female Artist 2__2	525,000	528	1,036	15,400	0.001004	0.00197	0.02928	Save-led
Female Artist 2__3	15,000	510	904	6,781	0.032903	0.058323	0.437484	Save-led
Female Artist 3__1	82,000	96	157	568	0.001172	0.001917	0.006936	Save-led
Female Artist 3__2	1,361,000	619	2,295	13,700	0.000455	0.001685	0.010061	Save-led
Female Artist 3__3	1,900,000	1,777	11,200	35,200	0.000935	0.005895	0.018526	Save-led
Female artist 4__1	1,800,000	2,097	10,400	40,900	0.001165	0.005778	0.022722	Save-led
Female artist 4__2	14,900,000	13,200	10,900	57,100	0.000886	0.000732	0.003832	Save-led
Female artist 4__3	245,400	411	1,425	3,450	0.001675	0.005807	0.014059	Save-led
Female Artist 5__1	1,600,000	1,827	222,900	29,300	0.001142	0.139313	0.018312	Share-led
Female Artist 5__2	1,000,000	1,120	15,118	21,100	0.00112	0.015118	0.0211	Save-led
Female Artist 5__3	821,900	992	11,273	15,800	0.001207	0.013716	0.019224	Save-led
Male Artist 1__1	6,100,000	10,100	52,900	127,800	0.001656	0.008672	0.020951	Save-led
Male Artist 1__2	1,300,000	1,000	2,600	8,800	0.000769	0.002	0.006769	Save-led
Male Artist 1__3	66,400	137	166	661	0.002063	0.0025	0.009955	Save-led
Male Artist 2__1	862,800	1,400	14,700	23,300	0.001623	0.017038	0.027005	Save-led
Male Artist 2__2	727,100	1,431	12,400	16,800	0.001968	0.017054	0.023105	Save-led
Male Artist 2__3	1,800,000	152	2,367	6,133	0.000084	0.001315	0.003407	Save-led
Male Artist 3__1	999,000	601	8,093	5,810	0.000602	0.008101	0.005816	Share-led
Male Artist 3__2	671,800	341	3,843	1,968	0.000508	0.00572	0.002929	Share-led
Male Artist 3__3	204,600	345	3,270	1,414	0.001686	0.015982	0.006911	Share-led

Male Artist 4__1	17,700,000	3,587	69,000	167,700	0.000203	0.003898	0.009475	Save-led
Male Artist 4__2	13,600,000	2,415	27,300	89,000	0.000178	0.002007	0.006544	Save-led
Male Artist 4__3	12,600,000	1,902	21,000	71,800	0.000151	0.001667	0.005698	Save-led
Male Artist 5__1	5,400,000	6,698	161,100	47,900	0.00124	0.029833	0.00887	Share-led
Male Artist 5__2	1,600,000	4,944	34,500	45,600	0.00309	0.021562	0.0285	Save-led
Male Artist 5__3	1,100,000	2,143	19,842	24,100	0.001948	0.018038	0.021909	Save-led