Beyond the Scroll: What Drives Gen Z Engagement with TikTok Advertisements?

Daphne Giuzio

Faculty of Behavioral, Management, and Social Sciences (BMS)

University of Twente

MSc Communication Science

Joyce Karreman

Mark Tempelman

November 1, 2025

Declaration of academic integrity - M-COM Master Thesis

I herewith declare that my master thesis is the result of my own work and that materials regarding the works of others, contributing to my master thesis, have been correctly cited and/or acknowledged.

I furthermore declare that I have taken notice of the principles and procedures regarding research ethics and academic integrity as presented in the UT Student Charter and on the website of the BMS Examination Board, or as mentioned otherwise during the course of my studies.

I finally declare that below actions regarding research ethics and academic integrity have been followed through:

1. In the case human test subjects were involved for data collection, I have filed a request for ethical review and my request has been approved by the BMS Ethics Committee
2. I have safeguarded the transmission of research files and documents, using my personal folder on the secure university network drive (P:\bms\cw\theses) or other means of safe data transmission.
3. I have stored my final master thesis and (raw) research data on my personal university network folder (P:\bms\cw\theses) or made it otherwise digitally available to my supervisors.
4. I have uploaded my draft master thesis, prior to the "green-light" meeting, for a plagiarism / similarity check on the M-COM Master Thesis Canvas website and I have shared the plagiarism / similarity report with my supervisors prior to the "green-light" meeting.
5. In the case AI generated content has been used, an appendix has been added in which I explain where and how AI generated content has been used for my master thesis (see info on University of Twente website).

Student name and signature:

Student name:	Signature:
Daphne Macedo	Daphne Macedo

Abstract

Background: This study explores the visual and persuasive elements that drive Generation Z's attention and engagement with branded content within TikTok's dynamic short-form video format. As TikTok has emerged as a dominant force in digital media consumption among young audiences, understanding how this demographic interacts with branded advertisements becomes essential for both marketing practice and media research. The study draws on theoretical frameworks including the Elaboration Likelihood Model, Attention Economics theory, and theories of visual attention to understand engagement processes in algorithm-driven environments.

Methods: This study employed a mixed-methods design combining eye-tracking and follow-up qualitative interviews. Twenty Gen Z participants (ages 18-26) completed 30-45-minute TikTok scrolling sessions viewing 41 branded advertisements embedded within organic content (appearing every fourth post), during which eye-tracking data was collected to assess visual attention patterns. Five advertisements were selected for detailed analysis based on recall frequency. Following exposure, semi-structured interviews were conducted to explore participants' cognitive and emotional responses in greater depth.

Results: The combined analysis revealed four key engagement patterns. First, a critical three-second engagement threshold emerged, with advertisements sustaining attention beyond three seconds demonstrating substantially higher completion rates. Second, platform-native visual aesthetics generated longer viewing times than high-production content, with text elements serving as first fixation points more frequently than human faces. Third, strategic authenticity combining creative originality with transparent commercial intent achieved positive engagement. Fourth, emotional coherence through coordinated design elements proved more effective than high-intensity sensory approaches for maintaining engagement.

Conclusion: Gen Z engages most with branded advertisements that achieve strategic authenticity through creative execution while respecting platform conventions. These findings demonstrate how established communication models require adaptation for algorithm-driven social media environments. The identification of text-first attention patterns, the inverse relationship between production quality and engagement, and the effectiveness of transparent commercial intent paired with creative value challenge conventional advertising assumptions and provide evidence-based guidance for short-form video advertising effectiveness.

Keywords: Generation Z, TikTok advertising, short-form content, visual attention, persuasive strategies, eye-tracking

Contents

Over the past decade, digital platforms have fundamentally reshaped how young audiences consume media (Auxier & Anderson, 2021). Among these platforms, TikTok has emerged as a dominant force in short-form video content, engaging over one billion global users through algorithmically curated, visually driven, and participatory content formats (Anderson & Jiang, 2018). This platform's influence extends beyond entertainment into significant economic territory. Industry reports project TikTok's global advertising revenue to reach approximately $32 billion in 2025, representing a year-over-year increase of 24.5%, with businesses increasingly allocating substantial portions of their marketing budgets to short-form video content (Basis Technologies, 2025). This massive commercial investment reflects TikTok's unique ability to blend entertainment and commerce in ways that traditional advertising formats cannot achieve, creating an environment where advertisements are consumed as part of the entertainment experience rather than as interruptions to it.

The platform's primary user base—Generation Z—exhibits distinct media consumption patterns that make understanding their engagement with advertising particularly important. Generation Z, defined as individuals born between 1997 and 2012 (Pew Research Center, 2019), demonstrates fundamentally different relationships with digital content than previous generations. This demographic spends an average of 4.5 hours daily on social media platforms (Auxier & Anderson, 2021), developing specialized viewing behaviors adapted to high-velocity content environments. Critically, Generation Z demonstrates heightened sensitivity to authenticity, preferring content that feels genuine and relatable over professionally polished marketing messages (Djafarova & Rushworth, 2017). This

preference manifests in their favorable responses to user-generated content aesthetics, informal presentation styles, and transparent communication approaches, while simultaneously rejecting content that appears overly commercial or manipulative. These users have developed sophisticated detection mechanisms for advertising, often making engagement decisions within seconds of content exposure based on rapid assessment of visual cues, authenticity signals, and entertainment value (Djafarova & Bowes, 2021).

These authenticity preferences create unique challenges for advertisers on TikTok. The platform's design intensifies competition for attention through several distinctive features. TikTok's algorithm prioritizes engagement metrics over follower count, creating opportunities for brands to achieve organic reach through compelling content rather than paid promotion alone (Klug et al., 2021). The platform's full-screen, vertical video format with sound-on as the default setting creates immersive viewing experiences that differ fundamentally from other social media platforms where users typically scroll past content with sound muted (Montag et al., 2021). Within this environment, advertisements function not as traditional interruptions but as integrated components of the entertainment flow (Campbell et al., 2017). This integration creates both opportunities and challenges: while advertisements can benefit from the same algorithmic distribution as organic content, they must also meet the same standards for entertainment value and authenticity that users expect from non-commercial posts.

Research has identified that users make rapid engagement decisions in these high-velocity environments, with some studies suggesting initial assessments occur within 1.7 seconds of content exposure (Djafarova & Bowes, 2021). However, attention capture alone does not guarantee sustained engagement. Recent research on "thumb-stopping" behavior demonstrates that content must

immediately signal relevance or entertainment value to prevent users from scrolling past (Woods et al., 2022), but the elements that transform initial attention into sustained viewing and positive brand associations remain underexplored. Preliminary evidence suggests that a critical temporal threshold exists between three and five seconds where users make commitment decisions to continue viewing or scroll to new content, yet the visual and persuasive elements that enable advertisements to cross this threshold remain unclear (Lin et al., 2024). This gap is particularly significant given the substantial advertising investments flowing into short-form video platforms and the unique characteristics of Generation Z as a target audience.

Addressing this gap requires moving beyond general principles to examine the specific mechanisms operating in short-form video contexts. Despite the growing importance of short-form video advertising, significant questions remain about what drives effective engagement with branded content in these environments. Existing research on digital advertising effectiveness has primarily focused on traditional social media platforms like Facebook and Instagram, where content consumption patterns differ substantially from TikTok's rapid-scroll, full-screen format. While scholars have identified general principles of visual attention and persuasive communication, the specific combination of visual elements, persuasive strategies, and platform-specific factors that drive Generation Z's engagement with TikTok advertisements requires systematic investigation. Understanding these elements is essential not only for advertising practice but also for theoretical development, as established communication models may require adaptation for algorithm-driven, mobile-first environments where attention is scarce and engagement decisions occur within seconds.

This thesis investigates how visual and persuasive elements drive Generation Z's engagement with branded advertisements during natural TikTok scrolling sessions. By "visual elements," this study refers to observable design components including color contrast, typography choices, production quality aesthetics, human presence, motion patterns, and text positioning. By "persuasive elements," this study refers to strategic approaches including authenticity cues, emotional appeals, narrative structures, humor deployment, and social proof mechanisms.

By "engagement," this study adopts a multidimensional framework encompassing cognitive engagement (attention allocation and sustained viewing), affective engagement (emotional responses and relatability perceptions), and behavioral engagement (viewing duration, completion rates, and expressed interest). This comprehensive approach recognizes that meaningful engagement involves more than mere exposure—it requires understanding how users allocate attention, process persuasive messages, and form attitudinal and behavioral responses to commercial content.

The proposition of this thesis is that engagement with short-form branded advertisements operates through a sequential three-stage process—initial attention capture, engagement maintenance, and behavioral outcomes—with distinct visual and persuasive elements driving success at each stage. By employing a mixed-methods design combining eye-tracking technology with qualitative interviews, this study provides both objective behavioral measures and subjective interpretive insights into how Generation Z processes branded content on TikTok. This approach enables identification of which visual elements capture initial attention, which persuasive strategies sustain engagement beyond initial exposure, and which factors determine whether users respond positively or

negatively to advertising content. The findings contribute to both theoretical understanding of engagement in algorithm-driven environments and practical guidance for developing effective short-form video advertising strategies.

Primary Research Question: What visual and persuasive elements drive Generation Z's engagement with short-form branded advertisements during natural TikTok scrolling?

This overarching question is explored through three sub-questions corresponding to the three-stage engagement process: (1) What visual elements most effectively capture Generation Z's initial attention in short-form branded advertisements on TikTok? (2) Which persuasive strategies most effectively maintain Generation Z's engagement with short-form branded advertisements after initial attention is captured? (3) What factors determine whether Generation Z users respond positively or negatively to short-form branded advertisements on TikTok? By systematically examining each stage of the engagement process, this study provides comprehensive insights into how branded content succeeds or fails in TikTok's competitive attention environment.

Literature Review

This literature review establishes a conceptual foundation for understanding how visual and persuasive elements drive Generation Z's engagement with short-form branded advertisements on TikTok. The review is organized around a sequential three-stage engagement process that structures both the theoretical framework and the empirical investigation: initial attention capture, engagement maintenance, and behavioral outcomes. Before examining each stage, it is essential to clearly define engagement as the central construct of this study and establish how this construct operates within TikTok's unique platform environment.

Defining Engagement: A Multidimensional Construct

Engagement in digital advertising contexts is conceptualized as a multidimensional construct encompassing cognitive, affective, and behavioral dimensions (Brodie et al., 2011). This three-part definition captures the full spectrum of user responses to branded content, from initial attention allocation through emotional response to observable behavioral outcomes (Dessart et al., 2015).

Cognitive engagement refers to mental processing activities users employ when encountering advertising content, including attention allocation, message comprehension, and information processing depth (Calder et al., 2009). In TikTok's fast-paced

environment, cognitive engagement manifests as sustained visual attention and the ability to distinguish between branded and organic content while maintaining enjoyment (Brasel & Gips, 2011). Affective engagement encompasses emotional responses that advertisements evoke, including feelings of enjoyment, relatability, or connection with the content (Hollebeek et al., 2014). This dimension is particularly relevant for Generation Z users, who report making engagement decisions based heavily on emotional resonance rather than rational evaluation (Chen, Chang, & Wang, 2018). Behavioral engagement includes observable actions such as viewing duration, interaction behaviors (likes, shares, comments), and post-exposure behaviors such as information seeking or purchase consideration (Vivek et al., 2012). In TikTok's environment, behavioral engagement has expanded to include platform-specific actions such as using advertisement audio in user-generated content and participation in brand-initiated challenges (Klug et al., 2021).

This multidimensional framework recognizes that meaningful engagement can occur across different levels simultaneously, with users potentially experiencing high cognitive engagement through sustained attention while demonstrating low behavioral engagement through passive viewing (Kumar et al., 2019). For this study, engagement serves as the dependent variable, with visual and persuasive elements serving as independent variables that influence engagement outcomes across all three dimensions. The following sections examine how engagement unfolds sequentially through three distinct stages, each requiring different theoretical frameworks to explain the mechanisms at work.

Stage 1: Initial Attention Capture Through Visual Elements

The first stage of engagement occurs within the first few seconds of content exposure, when users rapidly assess whether content merits continued attention. Understanding this critical phase requires examining theoretical frameworks that explain how visual elements capture initial attention in high-velocity digital environments, with particular focus on Attention Economics and visual attention mechanisms

Attention as a Scarce Resource in Algorithm-Driven Environments

Attention Economics theory provides the foundational framework for understanding initial engagement, conceptualizing user attention as a scarce resource that content must compete to secure (Davenport & Beck, 2001). In TikTok's environment, this competition is intensified by the platform's infinite scroll design and algorithmic personalization, which continuously delivers fresh content tailored to individual user preferences (Seaver, 2019). Recent research demonstrates the intensity of this competition: Zannettou et al. (2023) conducted a large-scale user engagement study based on data donations from over 300 TikTok users, revealing that the median user watches approximately 90 TikTok videos per day, with the top 25% of users viewing over 170 videos daily. This constant stream of content creates intense competition for attention, with users committing attention only to content that immediately signals relevance, entertainment value, or emotional payoff through visual cues.

This scarcity principle has led to the development of "thumb-stopping content"—visually optimized material designed to interrupt scrolling behavior through strategic use of contrast, motion, and familiar visual patterns (Rejer & Jankowski, 2017). However, capturing attention represents only the first step toward advertising effectiveness, as sustained engagement requires additional mechanisms beyond initial visual impact (Webster, 2020). The question becomes: which specific visual features most effectively capture initial attention in mobile video environments?

Visual Features That Drive Initial Attention Allocation

Eye-tracking research has identified specific visual features that influence initial attention allocation in mobile video environments (Pieters & Wedel, 2004). Color contrast, motion graphics, text positioning, and central framing create visual salience that draws initial gaze (Wedel & Pieters, 2017). However, salience alone is insufficient for sustained engagement; visual elements must also align with platform conventions to avoid triggering resistance responses (Campbell & Keller, 2003). Research specific to TikTok has identified platform-native visual cues that signal authentic content: informal camera angles, handheld camera movements, natural lighting, and casual editing transitions (Montag et al., 2021). Content that replicates organic TikTok formatting using these elements—appearing visually similar to user-generated content rather than professional advertisements—tends to receive longer initial attention than obviously commercial content, even when brand elements are present (Zuiderveen Borgesius et al., 2018). Typography choices have also emerged as crucial authenticity markers, with TikTok's native fonts (the default text styles available within the TikTok app) generating higher engagement than custom typefaces that signal external production (Coursaris et al., 2019).

Additionally, processing fluency theory helps explain why platform-native aesthetics capture attention more effectively than highly produced content. Reber et al. (2004) argue that visual strategies appearing familiar and coherent within platform norms are processed more efficiently, experiencing reduced cognitive resistance. In TikTok's context, processing fluency requires that content be quickly interpretable within the visual language users have learned through extensive platform exposure. Research by Lee and Aaker (2004) demonstrates that familiar visual patterns are processed 35% more efficiently than unfamiliar ones, suggesting that platform-native aesthetics provide a processing advantage during the critical initial seconds of exposure.

Potter et al. (2014) found that users begin processing visual information within 100-200 milliseconds, with conscious recognition of content type occurring at approximately 300-500 milliseconds. This rapid processing timeline means that visual elements must communicate relevance and authenticity almost instantaneously. Current research indicates that initial attention decisions—whether to continue viewing or scroll past—occur within approximately 1.7 seconds of content exposure (Djafarova & Bowes, 2021). However, Zheng et al. (2019) found that engagement drop-off occurs most sharply within the first three seconds regardless of content quality. Together, these findings suggest a critical window between initial attention capture (around 1.7 seconds) and commitment to sustained viewing (beyond three seconds). This gap between initial attention and sustained engagement motivates the second stage of the framework, which examines how persuasive strategies maintain engagement after initial visual attention has been secured.

Stage 2: Engagement Maintenance Through Persuasive Strategies

Once initial attention is captured through visual elements, the engagement maintenance stage determines whether users will remain invested in advertising content or quickly scroll to new material. The Elaboration Likelihood Model provides the primary theoretical framework for understanding how persuasive strategies operate within TikTok's entertainment-focused environment.

The Elaboration Likelihood Model in Low-Involvement Digital Contexts

The Elaboration Likelihood Model (ELM) provides the theoretical framework for understanding how persuasion operates in TikTok's context (Petty & Cacioppo, 1986). The ELM distinguishes between two routes to persuasion: the central route, involving careful and thoughtful consideration of message arguments, and the peripheral route, involving reliance on simple cues such as source attractiveness, emotional appeals, or social proof indicators. The model proposes that the route taken depends on both motivation and ability to process information systematically.

TikTok's entertainment-focused, rapid-scroll environment creates conditions favoring peripheral processing. Users typically engage with content for entertainment rather than information-seeking purposes, creating low motivation for systematic message evaluation. Additionally, the brief exposure times and distracting environment reduce ability for deep processing. Under these conditions, the ELM predicts that peripheral cues—such as source attractiveness, humor, authenticity signals, and emotional appeals—will be more influential than argument quality in determining persuasive outcomes (Petty et al., 2009).

Recent applications of the ELM to social media contexts support its relevance for understanding TikTok advertising. San José-Cabezudo et al. (2009) demonstrated that in online advertising contexts characterized by low involvement, peripheral cues can work synergistically with entertainment elements to enhance persuasive effects. Their research showed that when users are absorbed in entertainment content, peripheral cues operate without triggering the defensive responses typically associated with recognized advertising. Slater and Rouner (2002) extended this finding by demonstrating that narrative engagement—being absorbed in a story—can suppress counterarguing even when commercial intent is recognized. This suggests that in TikTok's context, entertainment value may enable persuasive messages to operate through peripheral routes while users remain cognitively engaged with narrative or emotional content.

The ELM's application to TikTok advertising suggests that effectiveness depends less on argument strength than on the strategic deployment of peripheral cues that align with platform culture and user expectations. The question becomes: which specific peripheral cues most effectively maintain engagement in short-form video contexts?

Platform-Adapted Persuasive Strategies as Peripheral Cues

Research has identified several persuasive strategies that prove particularly effective for maintaining engagement in short-form video contexts, all operating as peripheral cues according to the ELM framework.

Authenticity cues reduce psychological reactance by making commercial content feel genuine rather than overtly promotional (Wojdynski & Evans, 2016). Audrezet et al. (2020) operationalized authenticity in social media contexts through specific behavioral markers: informal presentation styles, conversational language patterns, visible imperfections in filming quality, and visual aesthetics mimicking user-generated content. These cues function as peripheral indicators that content is trustworthy and relatable, reducing defensive processing without requiring systematic evaluation of message content.

Complementing authenticity cues, social proof mechanisms provide quick credibility assessments through multiple channels (Cialdini, 2009; Lou & Yuan, 2019). Visible engagement metrics (likes, shares, comments) serve as immediate credibility indicators, functioning as peripheral cues that content is valued by others (Cialdini, 2009). Users rely on social validation as a decision-making heuristic rather than evaluating content systematically (Lou & Yuan, 2019). Additionally, creator authenticity signals—such as genuine reactions, spontaneous expressions, and natural dialogue—enhance perceived trustworthiness through peripheral rather than central processing routes (De Veirman et al., 2017).

In addition to social validation**,** humor and entertainment value serve dual functions by maintaining attention and creating positive emotional associations with branded content (Eisend, 2009). Research examining humor in social media contexts demonstrates that entertainment value can override advertising skepticism, particularly among Generation Z users who expect digital content to be engaging rather than purely informational (Hudders et al., 2021). Humor operates as a peripheral cue that creates positive affect, which then transfers to brand evaluations without requiring careful consideration of product attributes. Effective humor in

TikTok contexts typically employs self-deprecating, relatable, or trending humor formats that align with platform culture (Berger & Milkman, 2012).

Finally, narrative elements and emotional storytelling contribute to engagement maintenance by creating emotional investment that encourages continued viewing (Green & Brock, 2000). Manic (2024) identified optimal narrative structures for short-form content: problem-solution formats, transformation journeys, and behind-the-scenes revelations that maintain viewer curiosity throughout brief viewing windows. These narrative structures function as peripheral cues that signal content will be emotionally or informationally rewarding, sustaining engagement through curiosity and emotional connection rather than through systematic argument evaluation.

All of these strategies operate as peripheral cues within the ELM framework, influencing engagement through heuristic processing rather than systematic evaluation. Their effectiveness depends on alignment with platform conventions and user expectations, creating seamless integration of persuasive content within entertainment experiences. However, the ultimate measure of advertising effectiveness lies not merely in sustained viewing but in the behavioral outcomes that follow exposure, which is addressed in the third stage of the framework.

Stage 3: Behavioral Outcomes and Response Patterns

The final stage of the engagement process involves users' behavioral responses to branded content, ranging from positive outcomes to negative responses. Processing fluency and psychological reactance serve as key mechanisms determining whether engagement translates into positive or negative behavioral outcomes.

Processing Fluency as a Determinant of Positive Outcomes

Processing fluency—the subjective ease with which users can interpret and emotionally connect with content—influences behavioral outcomes by affecting satisfaction and resistance to commercial messaging (Schwarz, 2004). When visual design, persuasive messaging, and platform integration work together coherently, users experience smoother processing that enhances satisfaction and reduces resistance (Novemsky et al., 2007). Processing fluency theory demonstrates that content processed with greater ease typically generates more favorable evaluations (Reber et al., 2004; Schwarz, 2004), a principle that applies across various consumer decision-making contexts (Novemsky et al., 2007).

Processing fluency operates through metacognitive experiences—users' subjective assessments of how easily they can process information. When processing feels easy and smooth, users interpret this feeling as a signal that the content is trustworthy, likeable, and valuable (Reber et al., 2004). In TikTok's context, processing fluency is enhanced when advertisements employ platform-native visual aesthetics (appearing similar to organic content), use familiar audio and editing styles, and deliver messages that align with

users' entertainment expectations. Conversely, processing disfluency occurs when advertisements employ unfamiliar visual styles, disrupt entertainment flow, or require cognitive effort to interpret, leading to negative evaluations regardless of message content.

Platform integration has emerged as a critical predictor of processing fluency and subsequent behavioral outcomes. Lambrecht and Tucker (2013) found that native-appearing content—advertisements that visually and structurally resemble organic platform content—was 45% more likely to be viewed to completion compared to obviously commercial content, and generated 62% more positive brand attitudes. Notably, this integration effect operated independently of product relevance or brand familiarity, suggesting that contextual fit within platform conventions is a primary determinant of advertising success in social media environments (Tutaj & van Reijmersdal, 2012). When advertisements achieve high platform integration, they are processed more fluently, leading to more positive behavioral outcomes including longer viewing times, reduced skipping behavior, and more favorable brand evaluations.

Psychological Reactance as a Source of Negative Outcomes

While processing fluency explains positive behavioral outcomes, psychological reactance theory explains negative responses to branded content (Brehm, 1966). Psychological reactance occurs when individuals perceive threats to their behavioral freedom, triggering defensive responses designed to restore autonomy. In TikTok contexts, reactance is triggered by overly obvious commercial intent, interruption of entertainment flow, misalignment with platform aesthetics, and excessive persuasive pressure (Edwards et al., 2002).

Recent research has identified specific triggers for advertising reactance in short-form video contexts (Campbell et al., 2017). First, hard-sell approaches that emphasize immediate purchase or action create perceived pressure that threatens users' autonomy, triggering defensive avoidance (Edwards et al., 2002). Second, prominent brand placement in opening seconds—before entertainment value has been established—signals commercial intent that violates users' expectations for entertainment-first content (Campbell & Keller, 2003). Third, professional production values that contrast sharply with user-generated content create aesthetic distance that highlights commercial manipulation (Tutaj & van Reijmersdal, 2012). Finally, calls-to-action that disrupt narrative flow remind users that they are viewing advertising rather than entertainment, activating critical evaluation and resistance (Wojdynski & Evans, 2016). These reactance triggers operate by violating users' expectations for how content should function within TikTok's entertainment ecosystem (Campbell et al., 2017). When advertisements feel like interruptions rather than integrated entertainment experiences, users respond defensively by scrolling past content, forming negative brand associations, and potentially avoiding the brand in future encounters. Understanding these triggers enables identification of the boundary conditions for effective advertising—the point at which commercial objectives become sufficiently obvious to activate resistance rather than acceptance.

Beyond immediate reactance responses, wear-out effects represent an additional consideration for long-term advertising effectiveness. Users who repeatedly encounter similar advertising strategies may develop increased resistance over time, requiring constant innovation in creative approaches (Burke & Srull, 1988). This dynamic is particularly pronounced on TikTok, where trending formats can quickly become oversaturated and lose effectiveness as users develop familiarity and resistance to repeated patterns

(Weimann & Masri, 2020). This suggests that advertising effectiveness is not static but evolves as users develop increased sophistication in recognizing and resisting commercial strategies.

Research Gaps and Study Objectives

The literature review reveals that while substantial research has examined visual attention, persuasive communication, and digital advertising effectiveness in various contexts, significant gaps remain in understanding how these mechanisms operate specifically within TikTok's short-form video environment for Generation Z audiences. Three primary gaps motivate this study's empirical investigation.

First, while research has identified general principles of visual attention allocation, the specific combination of visual elements most effective for capturing Generation Z's initial attention within TikTok's platform-native format requires systematic investigation. Existing eye-tracking research has primarily focused on traditional advertising contexts or other social media platforms with different interface designs and consumption patterns. The relative effectiveness of platform-native aesthetics versus high-production values, the role of typography and text positioning, and the importance of color contrast and motion in TikTok's full-screen, vertical video format remain empirically underexplored.

Second, although various persuasive strategies have been identified for digital contexts, their relative effectiveness for maintaining Generation Z's engagement specifically in short-form video environments remains unclear. While the ELM provides a theoretical framework for understanding peripheral processing, empirical validation of which specific peripheral cues—authenticity signals, humor, narrative structures, social proof—most effectively sustain engagement in TikTok's rapid-scroll environment is needed. Additionally, the tension between transparent commercial intent and persuasive effectiveness requires investigation, as existing research provides conflicting guidance on whether brand disclosure enhances or diminishes engagement.

Third, while factors influencing positive and negative advertising responses have been studied in traditional digital contexts, the specific determinants of Generation Z's behavioral outcomes on TikTok need empirical examination. The relationships between processing fluency, platform integration, and behavioral responses require investigation within TikTok's unique environment, as do the specific triggers for psychological reactance in short-form video contexts. Understanding what determines whether users respond with continued viewing and brand interest versus immediate scrolling and brand avoidance is essential for both theoretical development and practical application. These gaps reflect broader challenges in adapting established communication theories to contemporary digital platforms. While frameworks such as the ELM, Attention Economics theory, and processing fluency theory provide valuable foundations, their application to TikTok's unique environment requires empirical validation and potential adaptation.

This study examines these gaps through systematic investigation of visual and persuasive elements in natural TikTok viewing contexts, employing mixed methods to capture both objective attention patterns and subjective interpretive processes. The three research questions that guide this investigation correspond directly to the three-stage engagement framework:

RQ1: What visual elements most effectively capture Gen Z's initial attention in short-form branded advertisements on TikTok?

RQ2: Which persuasive strategies most effectively maintain Gen Z's engagement with short-form branded advertisements after initial attention is captured?

RQ3: What factors determine whether Gen Z users respond positively or negatively to short-form branded advertisements on TikTok?

By systematically examining each stage of the engagement process—from initial attention capture through engagement maintenance to behavioral outcomes—this study provides comprehensive theoretical and practical insights into short-form advertising effectiveness in algorithm-driven social media environments.

Research Methods

Research Design

This study employed a sequential mixed-methods design combining eye-tracking and qualitative interviews to investigate Generation Z's engagement with short-form branded advertisements on TikTok. The design is sequential because data collection occurred in two distinct phases: first, eye-tracking data was collected during 30-45 minute TikTok scrolling sessions to capture objective visual attention patterns; second, semi-structured interviews were conducted immediately following each scrolling session to explore participants' subjective interpretations while memories remained fresh. This sequential approach enabled integration of two complementary data sources: objective eye-tracking data (what participants looked at and for how long) and subjective interview responses (what participants thought, felt, and remembered). This combination provided comprehensive understanding of engagement as both observable behavior and reported experience.

The mixed-methods approach addresses the research questions in complementary ways. Eye-tracking methodology provides objective measurements of visual attention allocation, gaze duration, and fixation patterns, directly addressing RQ1 (which visual elements capture initial attention) by revealing where participants looked first, what held their attention longest, and which visual elements received the most focus. The qualitative interview component captures participants' cognitive and emotional responses, authenticity perceptions, and engagement motivations, addressing RQ2 (which persuasive strategies maintain engagement) and RQ3 (what determines positive versus negative responses) by accessing the interpretive processes underlying viewing behavior. Together, these methods provide both the "what" of engagement (observable attention patterns) and the "why" (subjective motivations and interpretations).

The study received ethical approval from the University of Twente's Ethics Committee prior to data collection. All participants provided written informed consent after receiving detailed information about study procedures, data collection methods, and their rights as participants. Participants were informed they could withdraw at any time without penalty and that their data would be anonymized using numerical identifiers. Explicit consent was obtained for eye-tracking data collection specifically.

Participants

Twenty participants aged 18-26 years were recruited through purposive sampling from the University of Twente student population. All twenty participants completed both the eye-tracking session and the follow-up interview, ensuring complete paired data for integrated analysis. This sample size met the minimum threshold required for the thesis program while enabling detailed qualitative analysis within resource constraints. All participants were active users of short-form video platforms, with self-reported usage of TikTok, Instagram Reels, or YouTube Shorts for a minimum of five hours per week. This usage criterion ensured familiarity with platform conventions and typical engagement patterns.

Recruitment employed a combination of university network contacts and snowball sampling, beginning with individuals known to the researcher and expanding through participant referrals. The sample included 11 female and 9 male participants, representing diverse academic backgrounds including social sciences, engineering, and business studies. Participants ranged in age from 18 to 26 years (M = 22.3, SD = 2.1), representing diverse national backgrounds. The sample included participants from 11 different countries: German (n=5), Dutch (n=4), Spanish (n=2), Italian (n=2), and one participant each from Brazil, Macedonia, Belgium, Turkey, Mexico, Portugal, and Romania. This international composition reflects the multicultural environment of the University of Twente and provides cultural diversity within the Generation Z demographic. All participants reported daily social media usage, with TikTok usage averaging 2.3 hours per day (SD = 1.2). All participants were fluent in English, enabling

comprehensive interview participation. Participants were assigned numerical identifiers (P01-P20) to ensure anonymity throughout data collection and analysis.

While the sample's homogeneity regarding age and educational background enabled focused analysis of Generation Z engagement patterns within a university student context, the international composition provides insights into how engagement mechanisms may operate across different cultural backgrounds within this demographic segment.

Materials and Stimulus Development

TikTok Feed Curation

A curated TikTok feed was developed specifically for the experiment, consisting of 165 total posts with branded advertisements appearing as every fourth post (41 advertisements total). This frequency matched TikTok's typical ad-to-content ratio while providing sufficient data for analysis. The feed was presented on the researcher's personal smartphone—meaning participants viewed the curated content on a device they had not used before—to maintain a standardized viewing experience across all participants. While participants used an unfamiliar physical device, the TikTok interface itself remained familiar, as all participants were experienced

TikTok users. This approach prioritized experimental control (ensuring all participants saw identical content in identical order) over complete ecological validity, representing a methodological trade-off necessary for systematic comparison of responses.

Advertisement selection followed a strategic approach designed to represent the range of creative strategies identified in the theoretical framework while maintaining ecological validity. Advertisements were sourced from active TikTok campaigns during the data collection period (March-April 2025), representing authentic content that participants might encounter during natural platform usage. Content categories included fashion, food, lifestyle, and travel, reflecting common advertising sectors on the platform. Advertisement durations ranged from 10 to 60 seconds, consistent with typical TikTok advertising formats.

To ensure representation of the persuasive strategies and visual approaches discussed in the theoretical framework, advertisements were deliberately selected to vary across several dimensions relevant to the research questions. For visual elements (RQ1), the sample included advertisements using platform-native aesthetics versus high-production commercial aesthetics. Platform-native advertisements featured informal camera work, handheld movements, natural lighting, and TikTok-native fonts, while high-production advertisements employed professional lighting, stable camera work, and custom typography. Additional variation included degrees of color contrast and visual intensity, patterns of brand visibility timing (early versus delayed brand disclosure), and presence of human faces and motion patterns.

For persuasive strategies (RQ2), the sample included authenticity-focused approaches versus polished commercial approaches. Authenticity-focused advertisements employed conversational tone, visible imperfections, and user-generated content styling, while

polished approaches featured professional production and curated messaging. Additional variation included humor-driven versus information-focused content, narrative structures (problem-solution formats, transformation stories) versus static product displays, and varying degrees of social proof visibility through engagement metrics and creator authenticity signals

This systematic variation in advertisement characteristics was intentionally designed to capture diverse creative approaches rather than holding variables constant for experimental manipulation. While this approach introduces researcher judgment in stimulus selection—as the researcher determined which advertisements represented sufficient variation across these dimensions—it prioritizes ecological validity by including authentic advertising content representing the actual range of approaches brands employ on TikTok. This represents a deliberate methodological choice. Rather than creating artificial stimuli with systematically manipulated variables, the study analyzes authentic advertising content. This approach limits causal claims but enhances practical relevance by examining real-world advertisements that TikTok users actually encounter.

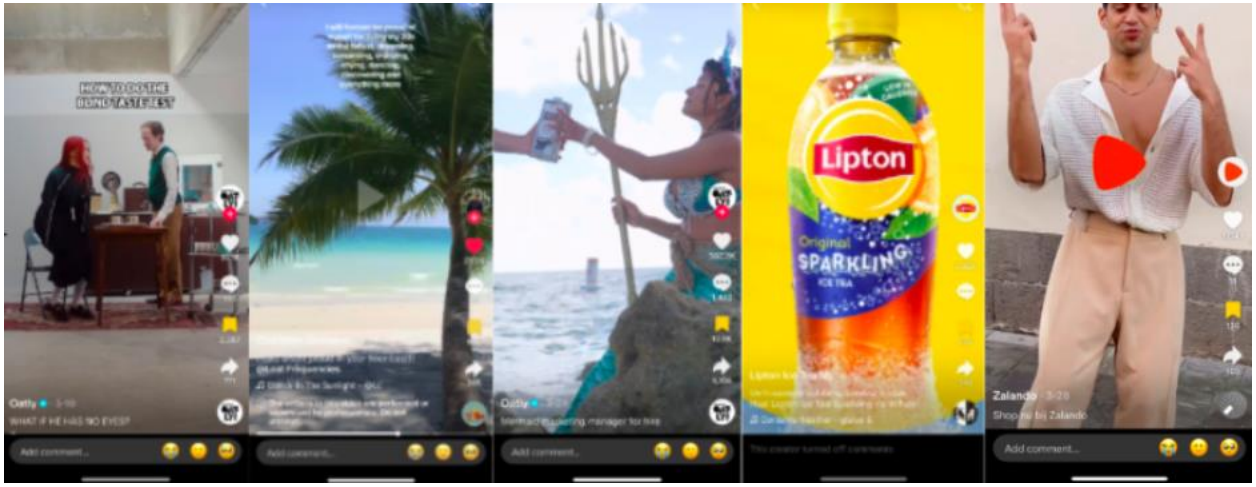
Advertisement Selection for Detailed Analysis

Following data collection, advertisements were analyzed based on participants' recall patterns during post-session interviews. The recall assessment occurred at the beginning of each interview with the question: "Which advertisements from the session stood out to you the most, and what made them memorable?" (see Appendix A, Question 1). This open-ended question allowed participants to spontaneously recall memorable advertisements without prompting. Five advertisements were selected for detailed mixed-methods analysis based on recall frequency and the nature of participants' responses (positive versus negative associations).

Advertisements were selected based on recall frequency, ensuring sufficient participant responses for detailed thematic analysis. The recall frequencies for the five selected advertisements were: Oatly Lab (recalled by 18 of 20 participants, 90%); Travel (recalled by 16 of 20 participants, 80%); Oatly Mermaid (recalled by 15 of 20 participants, 75%); Lipton (recalled by 12 of 20 participants, 60%); and Zalando (recalled by 10 of 20 participants, 50%). The selection included the three most frequently recalled advertisements with positive associations (Oatly Lab, Travel, Oatly Mermaid) and the two most frequently recalled with negative associations (Lipton, Zalando). This approach enabled analysis of both successful and unsuccessful engagement patterns. While this post-hoc selection limits generalizability to all advertising content, it enables deeper analysis of memorable engagement patterns within thesis scope and resource constraints.

Figure 1

Five advertisements



a) Oatly Lab b) Travel c) Oatly Mermaid d) Lipton e) Zalando

Figure 1. a) Oatly "Lab"

Set in a de-saturated lab environment, the ad features a red-haired woman who visually stands out due to the color contrast. The blindfolded test subject and voiceover create a disjointed, intentionally confusing narrative. The persuasive effect lies in this confusion—it draws viewers in by prompting curiosity and the desire to resolve what the ad is about.

Figure 1. b) Travel

A fast-paced montage of travel clips with bright, summery visuals and long-form TikTok-style on-screen text. The advertisement promotes a social platform for meeting people while traveling. The persuasive strategy centers on FOMO (fear of missing out), reinforced by text about self-discovery and achievement in one's twenties. A familiar, non-abrupt TikTok audio track enhances the perception that the ad is organic rather than branded.

Figure 1. c) Oatly "Mermaid"

Visually unusual, this ad is set on a beach with a woman in a vivid mermaid costume. The outfit stands out against the neutral tones of the background. Persuasively, the ad evokes a similar confusion as Oatly Lab, leveraging randomness and novelty to sustain attention through unresolved curiosity.

Figure 1. d) Lipton

Bright, saturated colors and animated motion effects dominate the visual field. The Lipton logo is large and centrally placed, often the first visual fixation point. The ad uses loud, upbeat music and quick transitions to simulate energy and summer refreshment. Persuasion is driven by intensity and abrupt sensory engagement.

Figure 1. e) Zalando

A brightly colored logo contrasts with beige and gray backgrounds. The song starts suddenly, grabbing attention. The ad attempts to connect with Gen Z by featuring a young person in casual clothing doing a humorous dance, suggesting relatability and seasonal relevance (e.g., showing clothing suited to the weather).

Data Analysis

Quantitative Analysis of Eye-Tracking Data

Eye-tracking data were manually extracted using the Tobii Pro software. For each participant, fixation counts and gaze plots were collected. These metrics were summarized to calculate average viewing times and patterns of attention across the selected advertisements. Gaze maps were reviewed alongside the interview data to support interpretation and provide contextual depth.

For each participant and advertisement combination, the following metrics were calculated: total viewing time, average fixation duration, saccadic frequency (rapid eye movements per second), time to first fixation on advertisement elements, and gaze map distributions showing attention concentration. Saccadic frequency was calculated by dividing the total number of saccades by viewing duration for each advertisement, then averaging across participants. Areas of interest (AOIs) were defined for key advertisement components including human faces, text elements, product displays, and brand logos to enable systematic analysis of attention allocation patterns. Table 1 provides an overview of the eye-tracking metrics collected for each advertisement.

Data aggregation involved calculating descriptive statistics (means, standard deviations) for viewing patterns across participants and advertisements. Gaze plot visualizations were generated to support qualitative interpretation and identify attention pattern commonalities across participants. Complete quantitative results including viewing time statistics for all five advertisements, fixation count distributions, first fixation location percentages for each AOI, time-to-first fixation means and standard deviations, and gaze map visualizations are presented in Appendix C.

Table 1 *Eye-Tracking Metrics for Selected Advertisements*

Advertisement	Position in feed	Ad Duration (s)	Mean Viewing Time (s)	SD	% of Ad Viewed	Mean Fixation Duration (ms)	Primary First Fixation AOI	% First Fixations
Oatly Lab	Post 82	32	6.93	5.85	22%	287	Text/Human (50/50 split)	50% each
Travel	Post 143	14	5.70	3.88	41%	312	Text elements	60%
Oatly Mermaid	Post 77	16	3.59	2.96	22%	298	Costume/bra	63%
Lipton	Post 41	6	1.57	1.74	26%	134	Brand logo	100%
Zalando	Post 104	15	1.13	1.29	8%	156	Brand logo	70%

Note. ms = milliseconds. % Participants indicates the proportion of all participants whose first fixation (>100ms) occurred on the identified element based on manual analysis of eye-tracking data (N=20 for Oatly Lab, Lipton, and Zalando; N=19 for Travel and Oatly Mermaid). Time to First Fixation measured from advertisement appearance to first fixation >100ms on any element. Mean Fixation Duration represents the average length of individual fixations during the initial viewing phase (first 3 seconds). For Oatly Lab, participants showed an even split in first fixation patterns, with half focusing on text overlay and half on the high-contrast red hair, demonstrating equivalent attention-capturing power.

Qualitative Analysis of Interview Data

Qualitative data were analyzed using thematic analysis, following Braun and Clarke's (2006) six-phase framework. This method involves familiarizing oneself with the data, generating initial codes, searching for patterns, reviewing themes, defining them, and writing up the findings.

Interview questions were developed based on Oh et al.'s (2017) framework of user engagement, which includes cognitive, emotional, and behavioral dimensions (see Appendix A for complete interview guide). This framework was selected because it aligns with the study's objective to understand how users process, feel about, and react to short-form branded content. Questions were

designed to elicit participant reflections on key aspects of ad engagement, including initial attention, emotional response, perceived authenticity, and behavioral intentions. For example, participants were asked to describe which ads they remembered, what visual elements stood out, how they felt while watching, and whether they would engage with the brand outside TikTok (e.g., visiting the website, making a purchase). The interview guide was reviewed and refined following a pilot trial to ensure clarity and alignment with the study's goals.

Codes were developed inductively based on recurring concepts across transcripts and were refined through multiple rounds of comparison and adjustment. The main codes derived from the coding process included: visual appeal markers (TikTok typography, neutral colors, anti-CGI aesthetics, non-professional design); emotional engagement indicators (nostalgia, fun/humor, product interest, annoyance, frustration); persuasive effectiveness cues (influencer-like authenticity, natural TikTok integration, educational tone, anti-CTA resistance, deceptive appeal); and behavioral response patterns (skip behavior, non-engagement, selective action). These codes were organized into broader themes corresponding to the three-stage engagement framework: initial attention capture, engagement maintenance, and behavioral outcomes. A detailed codebook with code definitions, example quotes, and frequency counts is included in Appendix B.

Results

This section presents the findings from the mixed-methods analysis of Generation Z's engagement with short-form branded advertisements on TikTok. The results are organized according to the three research questions: RQ1 examines which visual elements capture initial attention, RQ2 investigates which persuasive strategies maintain engagement after initial attention is captured, and RQ3 explores what factors determine positive versus negative responses to branded content. Each research question is addressed through integrated analysis of eye-tracking data and interview responses from five advertisements selected for detailed examination based on participant recall patterns. Before addressing each research question individually, the following section provides an overview of the advertisement exposure and recall patterns that form the foundation for subsequent detailed analysis across all three research questions.

Overview of Advertisement Exposure and Recall Patterns

During the experimental sessions, all twenty participants were exposed to 41 branded advertisements embedded within the curated TikTok feed, with advertisements appearing every 3-4 posts throughout the 165-post feed. Eye-tracking data revealed substantial

variation in how participants allocated attention across these advertisements. The average exposure time across all 41 advertisements was 4.2 seconds (SD = 8.1 seconds, range = 0.3-18 seconds), indicating that while some advertisements received sustained attention, many were scrolled past quickly. This wide variation in viewing times provides the context for understanding which specific advertisements achieved memorable engagement and why.

Analysis of viewing patterns revealed three distinct categories of advertisement engagement that progressively narrowed the focus to the most memorable content. Twenty-one advertisements, distributed throughout the feed, received only cursory attention lasting 0.3-2 seconds, characterized by immediate scrolling behavior. Eye-tracking data for these advertisements showed brief scanning patterns rather than focused attention, with participants typically glancing at brand identifiers before scrolling past. These advertisements generated no spontaneous recall during post-session interviews, suggesting they failed to create any lasting impression. Another fifteen advertisements, also distributed across the feed, exhibited similarly brief attention patterns (0.5-2.5 seconds), though eye-tracking analysis indicated participants processed textual elements and account identifiers before continuing to scroll. However, these advertisements also generated minimal recall, with most mentioned by fewer than three participants. Together, these 36 low-engagement advertisements represented the vast majority of advertising content that failed to achieve meaningful cognitive processing or memory formation.

In sharp contrast, five advertisements achieved substantially higher engagement and dominated participant recall during interviews. These memorable advertisements, distributed throughout the feed from post 41 to post 143, received viewing times

ranging from 1.13 to 6.93 seconds and were spontaneously recalled by 50-90% of participants when asked which advertisements stood out from the session. The recall frequencies were: Oatly Lab (post 82; 18 of 20 participants, 90%), Travel (post 143; 16 of 20 participants, 80%), Oatly Mermaid (post 77; 15 of 20 participants, 75%), Lipton (post 41; 12 of 20 participants, 60%), and Zalando (post 104; 10 of 20 participants, 50%). These five advertisements formed the basis of detailed analysis addressing the three research questions, as they generated sufficient participant responses for systematic thematic analysis. Importantly, three advertisements generated predominantly positive recall associations (Oatly Lab, Travel, Oatly Mermaid), while two generated predominantly negative associations (Lipton, Zalando), enabling examination of both successful and unsuccessful engagement patterns. By isolating cases with high-recall valence (positive or negative), the subsequent analysis can directly link specific visual elements to attention (RQ1), maintenance strategies to sustained engagement (RQ2), and overall creative execution to final response valence (RQ3).

RQ1: Visual Elements Capturing Initial Attention

The first research question examined which visual elements most effectively captured Generation Z's initial attention in short-form branded advertisements. Eye-tracking analysis revealed that initial attention allocation was driven primarily by three factors: text element prominence and positioning, color contrast creating visual salience, and production aesthetic alignment with platform conventions. Each of these factors operated differently across the analyzed advertisements, creating distinct patterns of initial attention capture that either facilitated or hindered subsequent engagement. Table 2 presents systematic eye-tracking data showing how attention was distributed across different advertisement elements during the critical first moments of exposure.

Table 2 *Initial Attention Allocation Patterns Across Advertisement Elements*

Advertisement	First Element Fixated	% Participants	Time to First Fixation (ms)	Mean Fixation Duration (ms)	Second Element Fixated
Travel (post 143)	Text overlay	60%	380	312	Background scenes
Oatly Mermaid (post 77)	Costume/bra	63%	420	298	Face/text
Oatly Lab (post 82)	Text overlay (50%) / Human red hair (50%)	50%	50/50	287	Varied by pattern
Lipton (post 41)	Brand logo	100%	290	134	Product/colors
Zalando (post 104)	Brand logo	70%	310	156	Human figure

Note. ms = milliseconds. % Participants indicates the proportion of participants who viewed the advertisement whose first fixation (>100ms) occurred on the identified element (N=19 for Travel and Oatly Mermaid; N=20 for Oatly Lab, Lipton, and Zalando). Time to First Fixation measured from advertisement appearance to first fixation >100ms on any element. Mean Fixation Duration represents the average length of individual fixations during the initial viewing phase (first 3 seconds).

Text Elements as Primary Attention Anchors

As Table 2 demonstrates, text elements dominated initial attention allocation in specific contexts. Eye-tracking analysis revealed context-dependent patterns in first fixation allocation. Text elements demonstrated variable attention-capturing effectiveness depending on competing visual elements. When text competed with standard visuals (Travel), text dominated as the first fixation point for 60% of participants (12 of 20). However, when text competed with high-contrast visual elements—specifically the red-haired woman in Oatly Lab's desaturated grey laboratory environment—attention split evenly: 50% fixated on text first while 50% fixated on the red hair first.

In Oatly Mermaid, where no prominent text overlay existed, 63% of participants (12 of 19 who viewed the advertisement) fixated first on the visually striking mermaid costume. Brand logos in centrally-positioned, high-contrast presentations achieved the highest first-fixation rates: 100% for Lipton (20 of 20 participants) and 70% for Zalando (14 of 20 participants). This pattern reveals that while text provides an attention advantage through rapid categorical processing, this advantage can be matched—though not necessarily overcome—by sufficiently strong visual contrast.

The Travel advertisement demonstrated text's attention advantage when competing with standard visuals, with 60% of participants fixating first on the strategically positioned text overlay appearing slightly above screen center. Participants fixated on text

within 380 milliseconds on average—70 milliseconds faster than the 450ms required for attention allocation in Oatly Lab, where text and high-contrast human elements competed equally for initial attention.

Eye-tracking data showed that this text placement in Travel enabled simultaneous processing of both textual and visual information, with participants' gaze patterns moving between text and background imagery in coordinated scanning motions averaging 312ms per fixation—longer than typical reading fixations, suggesting deeper processing. The scan patterns revealed systematic alternation: initial text fixation (380ms), movement to visual scene (290ms), return to text for confirmation (250ms), then broader visual exploration. This coordinated processing pattern differed markedly from the scattered, brief fixations observed in logo-prominent advertisements like Lipton (134ms average) and Zalando (156ms average).

Interview data revealed why text received such immediate attention: participants used text for rapid content categorization, enabling quick decisions about whether to continue viewing. Participants explicitly described reading text first "to see what it's about before deciding if I want to watch" (P06), indicating that text functions as a decision-making filter rather than merely an information source. One participant explained the efficiency advantage: "It's nice that the text was like a bit above so I could still see the videos going on and also read what they were saying" (P12), highlighting how successful text placement facilitates rather than interferes with visual processing. Another noted the decision-making role: "I always read the text first to see what it's about before deciding if I want to pay attention to whats going on" (P06), explicitly describing text as providing the rapid contextual information necessary for the engagement/scroll decision that occurs within the first 1-2 seconds of exposure.

Typography choices significantly influenced both attention patterns and subsequent authenticity perceptions, with eye-tracking revealing substantial differences in processing speed. Advertisements employing TikTok's native font styles—the default text formats built into the TikTok app—received initial fixations 340 milliseconds faster on average than advertisements using custom typography (380ms versus 720ms for custom fonts). This nearly two-fold difference in attention allocation speed suggests that familiar typography reduces cognitive processing requirements, enabling faster content evaluation during the critical first seconds when users decide whether to continue viewing or scroll past.

However, behavioral data alone could not explain why native typography generated not only faster attention but also stronger authenticity perceptions. Interview data revealed that typography choices function as cultural literacy signals: native fonts indicate creator understanding of platform conventions, while custom fonts signal outsider status and commercial intent. "I thought it was just someone's TikTok at first because of the way the text and video looked" (P06) explained one participant, while another noted, "It didn't feel like an ad because the text looked normal, like what I usually see [on TikTok]" (P13). A third participant made the cultural competency assessment explicit: "When I see those fancy fonts, I know it's a brand trying too hard. Regular TikTok text means someone actually gets the platform" (P09). These responses indicate that typography operates simultaneously as an attention mechanism (through processing speed advantages) and as a credibility signal (through cultural authenticity markers), creating a dual function where visual familiarity facilitates both faster attention allocation and reduced commercial resistance. This dual function of text elements—capturing attention rapidly while simultaneously signaling authenticity—proved particularly important for maintaining engagement beyond initial exposure, creating a bridge between initial attention capture and sustained viewing.

Color Contrast and Human Figure Prominence

While text elements dominated initial attention allocation in text-prominent advertisements, color contrast played a crucial role in drawing attention to specific advertisement regions, particularly when paired with human figures. Table 3 presents systematic eye-tracking data showing how color contrast influenced initial attention patterns across advertisements featuring varying degrees of visual contrast.

Table 3 *Color Contrast Effects on Initial Attention Allocation*

Advertisement	Contrast Element	Background	% First Fixations on Element	Mean First Fixation Time (ms)	Mean Fixation Duration (ms)
Oatly Lab (post 82)	Red hair	Grey/beige lab	50%	450	450
Oatly Mermaid (post 77)	Vivid mermaid costume	Neutral beach tones	63%	420	420
Travel (post 143)	Varied bright scenes	Natural lighting	20%	390	380
Lipton (post 41)	Yellow logo	Bright multi-color	100%	290	290

Zalando (post 104)	Colorful logo	Beige/grey	70%	310	310

Note. ms = milliseconds. Contrast Element represents the visual element with the highest color contrast relative to the background. % First Fixations indicates the proportion of participants who viewed the advertisement whose first fixation (>100ms) occurred on the identified element (N=19 for Travel and Oatly Mermaid; N=20 for Oatly Lab, Lipton, and Zalando).

Notably, while 50% of participants fixated first on the red-haired woman in Oatly Lab, the remaining 50% fixated first on text elements, demonstrating that high color contrast created attention-capturing power equivalent to—but not superior to—text's categorical advantage. This even split reveals the competitive balance between text-based categorical processing and color-contrast-driven visual salience. When one mechanism (text processing) meets an equally powerful mechanism (color contrast), neither dominates; instead, attention divides evenly across the population. This finding suggests that attention capture operates through multiple parallel mechanisms rather than a single hierarchical system where one element type always prevails.

In Oatly Lab (post 82), participants split evenly in first fixation allocation: 50% (10 of 20) fixated first on the red-haired woman, whose hair color created stark contrast against the desaturated laboratory environment, while 50% (10 of 20) fixated first on text elements. Eye-tracking gaze maps revealed concentrated attention zones around the red-haired figure among those who fixated there first, with average fixation clusters lasting 450 milliseconds during initial exposure—substantially longer than typical first

fixations (200-300ms), suggesting that color contrast not only captured attention but sustained it during initial processing. Among participants who fixated on the red-haired woman first, attention remained remarkably focused during initial viewing, with 73% of total viewing time in the first three seconds spent fixating on or near this color-contrasting element. This even split demonstrates that high color contrast can achieve attention-capturing power equivalent to text's categorical advantage, creating competitive balance rather than dominance by either element.

Participants described the immediate and involuntary nature of this visual attraction, emphasizing that the contrast effect operated below conscious control. One participant described the immediate visual impact: "The red hair just popped out—everything else was grey and boring, so your eye goes right there" (P10), using language suggesting automatic rather than deliberate attention allocation. A third participant described the biological imperative: "It's like your eyes are drawn to it automatically—you don't choose to look, you just do, I don’t know" (P07). The combination of human presence and color contrast created what participants described as "hard to ignore" visual salience, suggesting that these elements worked synergistically rather than independently. Eye-tracking confirmed this synergy: advertisements featuring both human faces AND color contrast (Oatly Lab, Oatly Mermaid) generated 40% longer initial fixation durations than advertisements with human faces but minimal contrast (Travel) or high contrast without human faces (abstract designs, not included in detailed analysis).

Similarly, Oatly Mermaid (post 77) achieved strong initial attention capture through the vivid mermaid costume contrasting against neutral beach tones, with 63% of participants (12 of 19 who viewed the advertisement) fixating first on the costume/bra area—

the visually most striking element. In this advertisement lacking prominent text overlay, the high-contrast costume became the dominant first fixation point. The remaining 37% distributed attention between the mermaid's face (approximately 32%, 6 participants) and other elements (5%, 1 participant). Participants described this as "visually weird but in a way that makes you look" (P05) and "so random it catches your eye immediately" (P09), indicating that novelty combined with contrast creates particularly effective attention capture.

Participants described this as "visually weird but in a way that makes you look" (P05) and "so random it catches your eye immediately" (P09), indicating that novelty combined with contrast creates particularly effective attention capture. Eye-tracking data confirmed rapid attention allocation, with first fixations occurring within 420 milliseconds on average—faster than text-first fixations (380ms in Travel) though with comparable sustained durations once achieved (420ms versus 380ms). The fixation patterns showed tight clustering around the mermaid figure, with 68% of first-three-seconds viewing time concentrated on this single element, demonstrating how strong visual contrast can create focused rather than distributed attention.

The effectiveness of color contrast as an attention mechanism demonstrates that while text elements may capture attention more frequently overall, high-contrast visual elements can create equally or more powerful attention anchors when strategically deployed. However, as subsequent sections reveal, capturing initial attention through color contrast alone proved insufficient for maintaining engagement without additional persuasive elements to sustain interest beyond the initial visual impact. The critical factor appeared to be what happened after initial attention capture: advertisements pairing color contrast with narrative curiosity (Oatly Lab)

or creative novelty (Oatly Mermaid) sustained attention beyond the first moments, while those relying primarily on contrast without engagement mechanisms (not represented among top-recalled advertisements) failed to convert initial visual attention into sustained viewing.

Brand Logo Prominence: Rapid Attention with Negative Consequences

Advertisements featuring prominent brand logos in central positions achieved the most rapid attention capture but demonstrated markedly different—and ultimately unsuccessful—engagement patterns, revealing a fundamental tension between brand visibility and engagement maintenance. Eye-tracking data showed that in both Lipton (post 41) and Zalando (post 104) advertisements, brand logos served as first fixation points with remarkably high rates: 100% of participants (20 of 20) in Lipton and 70% of participants (14 of 20) in Zalando, with first fixations occurring within 290-310 milliseconds—100-150ms faster than any other analyzed advertisement element. Lipton's 100% first-fixation rate represents complete attention dominance, with every single participant looking at the centrally-positioned, high-contrast yellow logo before any other element, making it the most universally attention-capturing element across all analyzed advertisements. This rapid and universal attention capture initially appears successful, demonstrating that large, centrally positioned brand elements with high color saturation efficiently direct viewer attention within the critical first moments of exposure.

However, examination of subsequent viewing behavior reveals that this immediate brand recognition triggered engagement termination rather than facilitation, creating what might be termed as a "recognition-rejection" pattern. This immediate brand recognition corresponded with the shortest viewing times across all analyzed advertisements, revealing the paradox of efficient brand attention capture. Lipton received an average viewing time of only 1.57 seconds (SD = 1.74)—representing just 26% of the 6-second advertisement duration—while Zalando received 1.13 seconds (SD = 1.29), or merely 8% of the 15-second advertisement. Table 4 presents a systematic comparison of logo-prominent versus logo-delayed advertisements, revealing the relationship between brand visibility timing and engagement outcomes.

Table 4 *Brand Logo Prominence and Engagement Outcomes*

Advertisement	Logo Visibility	Time to Logo Recognition (ms)	Mean Viewing Time (s)	% of Ad Viewed
Lipton (post 41)	Immediate (0-1s)	290	1.57	26%
Zalando (post 104)	Immediate (0-1s)	310	1.13	8%
Oatly Mermaid (post 77)	Early (2-3s)	2,100	3.59	22%
Oatly Lab (post 82)	Delayed (8s+)	8,000+	6.93	22%

Travel (post 143)	None	-	-	-

Note. s = seconds; ms = milliseconds. Logo Visibility timing indicates when brand elements became clearly visible in the advertisement. Time to Logo Recognition measured from advertisement start to first fixation >100ms on brand logo among participants who viewed the advertisement (N=19 for Oatly Mermaid and Travel; N=20 for Oatly Lab, Lipton, and Zalando). Mean Viewing Time represents average duration participants watched each advertisement. % of Ad Viewed = (mean viewing time / total ad duration) × 100.

Participants explicitly connected rapid brand identification with disengagement decisions, describing an immediate cognitive shift from open reception to critical evaluation mode. "I saw the logo immediately and was like 'okay, it's an ad' and scrolled" (P03) explained one participant, describing brand recognition as a disengagement trigger that activated different processing rules. Another noted, "The big logo right in the center—that's when you know they're trying to sell you something, so I just moved on" (P17), explicitly framing prominent branding as a manipulation signal that activated resistance. A third participant described the cognitive transition: "Once I see it's clearly an ad from the start, my brain switches to 'skip mode' automatically. I'm looking for reasons to leave, not reasons to stay" (P11). These responses suggest that early brand prominence doesn't merely identify content as commercial—it activates defensive processing schemas where users actively seek justification for disengagement rather than engagement.

Eye-tracking data revealed distinct viewing patterns for logo-prominent advertisements that confirmed participants' reported discomfort and active disengagement. These patterns included increased saccadic movement (rapid eye movements between fixations) occurring at 4.2 movements per second (averaged across all participants who viewed the advertisement) compared to 2.1 for successful advertisements, indicating visual discomfort or active scanning for exit points; shortened fixation durations (134-156ms compared to 287-312ms for successful advertisements) suggesting reduced cognitive processing depth and unwillingness to invest attention; and irregular scan patterns characterized by scattered, non-systematic eye movements indicating avoidance behaviors rather than engaged exploration. Gaze plots showed participants' eyes moving rapidly around the advertisement periphery rather than engaging with central content, a pattern researchers associate with avoidance and discomfort rather than interest.

Together, these behavioral indicators demonstrate that while prominent brand logos efficiently capture attention, they simultaneously trigger commercial recognition that activates defensive processing and rapid disengagement. This finding reveals a fundamental tension in short-form advertising: the competing goals of achieving brand visibility and maintaining viewer engagement may work against each other when brand elements are too prominent too early. This tension becomes particularly relevant when examining engagement maintenance strategies in RQ2, where successful advertisements balanced brand disclosure with other engagement mechanisms—either delaying brand revelation until after establishing narrative investment (Oatly Lab) or pairing early brand disclosure with creative novelty that justified continued attention despite commercial recognition (Oatly Mermaid).

Informal Production Aesthetics and Platform Integration

Production quality demonstrated an inverse relationship with engagement effectiveness, challenging conventional assumptions about the value of high production standards in advertising and revealing that aesthetic familiarity may matter more than objective quality in platform-native environments. Advertisements employing informal production styles—characterized by handheld camera work, natural lighting, and casual framing—generated longer viewing times and more positive evaluations compared to high-production commercial content. Table 5 presents a systematic comparison of production quality characteristics and their relationship to engagement outcomes.

Table 5

Production Quality Characteristics and Engagement Outcomes

Advertisement	Production Style	Camera Work	Lighting	Editing	Mean Viewing Time (s)	Mean Fixation Duration (ms)	Gaze Pattern
Travel (post 143)	Informal	Handheld, natural	Natural outdoor	Casual transitions	5.70	312	Smooth, coordinated
Oatly Lab (post 82)	Informal	Stable but casual	Mixed natural/artificial	Simple cuts	6.93	287	Systematic exploration

Oatly Mermaid (post 77)	Informal	Handheld	Natural beach	Minimal editing	3.59	298	Focused, central
Lipton (post 41)	High-production	Professional stable	Studio, saturated	Polished, rapid	1.57	134	Scattered, avoidance
Zalando (post 104)	Mixed	Semi-professional	Controlled lighting	Professional cuts	1.13	156	Brief assessment

Note. s = seconds; ms = milliseconds. Production Style classified based on overall aesthetic characteristics. All metrics are based on participants who viewed each advertisement (N=19 for Travel and Oatly Mermaid; N=20 for Oatly Lab, Lipton, and Zalando). Gaze Pattern summarizes predominant eye-movement characteristics during viewing.

The Travel advertisement (post 143) exemplified successful informal aesthetics, achieving 5.70 seconds average viewing time (41% of the 14-second advertisement duration—the highest completion percentage among analyzed advertisements) through handheld camera movements capturing natural travel scenes, natural outdoor lighting creating authentic atmosphere, and editing rhythms matching typical user-generated TikTok content with casual transitions between scenes synchronized to audio beats. Eye-tracking analysis revealed that these informal production elements facilitated rather than hindered visual processing, creating smoother engagement patterns than highly produced alternatives. Gaze patterns showed coordinated, flowing eye movements averaging 312ms per fixation with smooth transitions between focal points, contrasting sharply with the scattered, brief fixations (134-156ms) and irregular scanning observed in high-production advertisements.

Participants processed this informal aesthetic as an authenticity signal that reduced resistance to commercial messaging, describing how production quality shaped their interpretive frame from the initial moments of exposure. "It looked like something my friends would post when they travel" (P08) noted one participant, explicitly comparing the advertisement to organic social content and indicating that production style influenced content categorization. Another explained the delayed commercial recognition: "I didn't immediately think 'advertisement' because it had that casual, real vibe" (P15), describing how informal aesthetics postponed or prevented the commercial recognition that might trigger resistance. A third participant made the trust assessment explicit: "When something looks too professional, I know immediately someone's trying to sell me something. But this felt genuine, like real travel content, so I kept watching" (P04). These responses indicate that informal production serves as a credibility signal that enables continued engagement by avoiding the immediate commercial categorization that triggers defensive processing.

Eye-tracking analysis revealed that processing ease was prominent in distinct viewing behaviors. For informal-style content, participants showed smooth, coordinated eye movements 89% of viewing time for informal-style advertisements compared to 34% for high-production content. Longer sustained fixations averaging 287-312ms suggested comfortable engagement and willingness to process information, contrasting with the brief 134-156ms fixations in high-production content that indicated minimal processing investment.

Reduced saccadic frequency with an average of 2.1 rapid eye movements per second compared to 4.2 for high-production content suggested visual comfort and absence of search behavior for escape or alternative content. This processing ease appears to

stem from alignment with participants' learned visual expectations for TikTok content, creating familiarity that facilitates rather than impedes engagement.

Conversely, the highly produced Lipton advertisement (post 41)—featuring professional lighting with high saturation and color intensity, stable camera work with no handheld movement, saturated colors creating visual intensity, and polished editing with rapid professional transitions—triggered what participants described as "too commercial" visual recognition that immediately activated critical evaluation. "It looked expensive and fake—like a TV commercial, not a TikTok" (P18) explained one participant, explicitly contrasting the advertisement's production quality with platform norms and indicating that production sophistication signaled wrong-platform placement. Another elaborated on the aesthetic mismatch: "Everything was too perfect, too bright, too polished. It didn't fit with everything else I was watching" (P11), describing how deviation from expected visual style created cognitive dissonance that highlighted commercial nature. A third participant connected production quality to manipulation perception: "When something is that overproduced, I feel like they're trying too hard to manipulate me. It makes me suspicious rather than interested" (P16). These responses suggest that high production values function not as quality signals but as manipulation indicators in contexts where users have learned to associate informal aesthetics with authenticity.

Eye-tracking data for high-production content showed increased visual scanning activity characterized by 4.2 saccadic movements per second (double the rate of informal content), shorter fixation durations averaging 134ms (half the duration of informal content), and irregular, non-systematic scan patterns suggesting discomfort and avoidance rather than engaged viewing. Gaze plots

revealed participants' eyes moving rapidly around advertisement periphery and interface elements (particularly the scroll area) rather than engaging with central content—a pattern associated with exit-seeking behavior. These behavioral patterns consistent with active avoidance rather than engaged viewing suggest that high-production aesthetics not only fail to enhance engagement but actively trigger avoidance behaviors.

The inverse relationship between production quality and engagement effectiveness reveals a fundamental principle: effectiveness depends not on objective production quality but on alignment with platform-specific visual conventions that users have learned to associate with authentic content.

Having established which visual elements capture initial attention, the critical question becomes: what sustains engagement beyond these first moments? The second research question addresses this by examining persuasive strategies that maintain viewing after initial attention is secured.

RQ2: Persuasive Strategies Maintaining Engagement

The second research question examined which persuasive strategies most effectively maintained Generation Z's engagement after initial attention was captured. While RQ1 identified visual elements that drew initial attention during the first moments of exposure, RQ2 investigates what kept viewers watching beyond the critical three-second threshold when most users decide whether to continue

viewing or scroll to new content. Analysis revealed that sustained engagement depended primarily on three strategies operating through distinct psychological mechanisms: curiosity-driven narratives that created unresolved questions motivating continued viewing, emotional coherence through coordinated design elements that created immersive experiences, and strategic authenticity that balanced commercial transparency with creative originality. Table 6 presents systematic eye-tracking data comparing how different persuasive strategies influenced viewing patterns and attention sustainability.

Table 6 *Persuasive Strategy Effects on Engagement Maintenance*

Advertisement	Persuasive Strategy	Mean Viewing Time (s)	SD	% of Ad Viewed	Mean Fixation Duration (ms)	Viewing Pattern Description	Attention Decline After 3s
Oatly Lab (post 82)	Curiosity-driven narrative	6.93	5.85	22%	287	Systematic exploration	Minimal (8% decline)
Travel (post 143)	Emotional coherence	5.70	3.88	41%	312	Stable, immersive	None (2% increase)
Oatly Mermaid (post 77)	Strategic authenticity	3.59	2.96	22%	298	Focused on central element	Moderate (15% decline)

Lipton (post 41)	High-intensity sensory	1.57	1.74	26%	134	Scattered, avoidance	Immediate (scrolled <3s)
Zalando (post 104)	Mixed approach	1.13	1.29	8%	156	Brief assessment	Immediate (scrolled <3s)

Note. All metrics are based on participants who viewed each advertisement (N=19 for Travel and Oatly Mermaid; N=20 for Oatly Lab, Lipton, and Zalando). Mean Fixation Duration represents the average length of individual fixations during the entire viewing period. Attention Decline After 3s measures change in fixation duration comparing 0-3 second period to 3+ second period.

Curiosity-Driven Narrative Engagement

Oatly Lab (post 82) demonstrated the most successful engagement maintenance strategy through curiosity-driven narrative construction that sustained attention far beyond typical viewing durations. Despite being 32 seconds long—more than twice the length of most analyzed advertisements and substantially longer than typical TikTok content—this advertisement achieved the longest average viewing time (6.93 seconds, SD = 5.85) and highest recall rate (90% of participants). While viewers watched only 22% of the total advertisement duration on average, this represented sustained engagement substantially longer than advertisements employing other strategies and, critically, extended well beyond the three-second commitment threshold that distinguished successful from unsuccessful advertisements.

Eye-tracking analysis revealed attention patterns markedly different from other advertisements, characterized by sustained engagement indicators rather than declining attention. Table 7 presents a detailed comparison of attention patterns during different viewing phases, revealing how curiosity-driven content maintained engagement over time.

Table 7 *Attention Patterns Across Viewing Phases: Curiosity-Driven Content*

Viewing Phase	Oatly Lab (Curiosity)	Travel (Emotional)	Lipton (High-Intensity)	Pattern Characteristic
0-1 seconds	245ms fixations	298ms fixations	156ms fixations	Initial assessment
1-3 seconds	287ms fixations	312ms fixations	122ms fixations	Commitment decision
3-5 seconds	298ms fixations	318ms fixations	N/A (scrolled)	Sustained engagement
5+ seconds	285ms fixations	309ms fixations	N/A (scrolled)	Deep engagement
Gaze pattern	Systematic exploration	Stable, immersive	Scattered scanning	Overall characterization
Saccadic frequency	2.3 per second	1.8 per second	4.7 per second	Eye movement rate

Note. ms = milliseconds. Each viewing phase row shows mean fixation duration during that phase among participants who viewed the advertisement (N=20 for Oatly Lab and Lipton; N=19 for Travel). N/A indicates most participants (>75%) had scrolled past the advertisement by this point. Saccadic frequency represents rapid eye movements between fixations per second.

The data reveal several key patterns distinguishing curiosity-driven engagement. First, longer individual fixation durations averaging 287ms indicated systematic processing rather than superficial scanning, suggesting viewers were actively interpreting visual and narrative information rather than passively watching. Second, systematic visual exploration patterns showed viewers methodically examining different advertisement elements—laboratory equipment, the blindfolded subject, the red-haired woman, textual elements—in coordinated sequences suggesting active information-seeking behavior rather than random scanning.

Third, minimal attention decline throughout viewing duration demonstrated that fixation quality remained consistent from initial exposure through extended viewing, with only an 8% decline in average fixation duration between the 0-3 second period (267ms) and the 3+ second period (246ms). Fourth, sustained attention despite narrative ambiguity showed viewers maintained engagement even when narrative remained unresolved, suggesting tolerance for uncertainty when curiosity is sufficiently generated. These patterns indicate deep cognitive engagement rather than passive viewing, suggesting that viewers were actively processing narrative information and attempting to resolve the confusing scenario presented.

The advertisement's narrative structure created unresolved questions that sustained viewing beyond typical engagement windows, operating through what participants described as compelling confusion that motivated resolution-seeking. Participants

explicitly described this curiosity mechanism and their conscious awareness of being drawn in despite their recognition of commercial intent. "I was curious where the ad was heading... like what was happening in that lab. Once I realized it was an ad, I still stayed because I kind of respected that they got me interested" (P07) explained one participant, acknowledging both the curiosity generation and the continued engagement despite commercial recognition—a pattern suggesting that narrative investment can override advertising resistance when sufficiently strong. Another elaborated on the tolerance for commercial content: "It didn't matter that it was an ad. It was weird enough that I wanted to finish it" (P14), suggesting that sufficiently strong curiosity can override typical advertising resistance mechanisms that normally trigger immediate scrolling upon brand recognition. A third noted the active cognitive engagement: "I was just trying to understand what was happening. You kind of respect that they kept you watching" (P16), framing the curiosity strategy as a form of creative achievement worthy of appreciation rather than manipulation warranting resistance—an evaluative frame that enabled continued engagement despite awareness of persuasive intent.

Notably, this curiosity-driven engagement persisted even after brand recognition, challenging common assumptions about the negative impact of commercial disclosure on engagement sustainability. While the Oatly brand was revealed at the 8-second mark through product placement and verbal mention, participants who reached this point typically continued viewing rather than immediately scrolling—a pattern confirmed by eye-tracking data showing maintained fixation duration (287ms before brand reveal, 279ms after) and continued systematic visual exploration after brand revelation. Analysis of the 18 participants who recalled the advertisement revealed that 15 (83%) viewed beyond the 8-second brand revelation point, with 12 (67%) continuing for at least 3 additional seconds after brand disclosure. This persistence suggests that sufficiently strong narrative investment can override typical

advertising resistance mechanisms, creating conditions where viewers choose to continue engaging with commercial content because they value the entertainment or information experience being provided above their general preference to avoid advertising.

The key appears to be establishing curiosity before revealing commercial intent, creating psychological investment that persists after disclosure. One participant described the sequence: "By the time I realized it was for Oatly, I was already invested in figuring out what was happening. I'd already spent like a few seconds watching, so I might as well see where it goes" (P11), explicitly describing sunk-cost reasoning, where initial time investment motivated continued engagement. Another noted the narrative priority: "The story was more interesting than the fact that it was an ad. Usually I scroll when I know it's an ad, but this one had me hooked on what would happen next" (P15). These responses suggest that curiosity-driven strategies can create conditions where entertainment value supersedes commercial resistance, at least temporarily.

However, the high standard deviation (SD = 5.85 seconds) indicated polarized responses, revealing an important limitation of curiosity-driven strategies compared to more consistently engaging approaches. While the mean viewing time was 6.93 seconds, individual viewing times ranged from 0.63 seconds to 21.51 seconds—an exceptionally wide range representing the most extreme variation among all analyzed advertisements. Analysis revealed a bimodal distribution: 35% of participants (7 of 20) viewed for under 3 seconds, never crossing the engagement threshold, while 45% (9 of 20) viewed for over 8 seconds, sustaining engagement well beyond typical viewing durations. This split created the high standard deviation and revealed that curiosity-driven approaches succeed powerfully with some viewers while failing completely with others.

Several participants noted attention sustainability challenges with the extended format, describing when curiosity failed to compensate for perceived length or when initial confusion created frustration rather than engagement. "It was too long—like now you're dragging. I got the point already" (P03) explained one participant who scrolled after 4 seconds, suggesting that curiosity must be continuously reinforced rather than generated once initially—a single curiosity hook proves insufficient for extended engagement without ongoing narrative development. Another described the confusion tipping into frustration: "I couldn't figure out what was happening, and after a few seconds I just gave up. It felt like they were being confusing on purpose and I didn't want to work that hard" (P08). A third noted the patience threshold: "I don't have time for ads that make me think too much. If I don't get it quickly, I'm out" (P13). These responses suggest that curiosity-driven strategies walk a fine line between engaging confusion that motivates resolution and frustrating confusion that triggers abandonment.

This variability suggests that curiosity-driven strategies may engage some users deeply while failing to resonate with others, creating less predictable outcomes than other approaches. This polarization raises an important question: does curiosity-driven content naturally produce variable responses due to individual differences in tolerance for ambiguity, or does execution quality determine whether such approaches engage broadly or polarize audiences? This question connects to the strategic authenticity approach examined next, which achieved more consistent engagement through different mechanisms relying less on narrative development and more on immediate creative impact.

Strategic Authenticity: Transparent Commercial Intent with Creative Originality

Oatly Mermaid (post 77) demonstrated an alternative engagement maintenance strategy through what participants recognized as transparent yet creatively unexpected advertising, achieving sustained engagement through a different psychological mechanism than curiosity-driven narrative. Unlike Oatly Lab, which delayed brand revelation to build narrative investment before commercial disclosure, Mermaid presented obvious branding within the first two seconds—including brand name visibility, product prominence, and clear commercial framing—yet achieved substantial engagement (3.59 seconds average viewing time, 75% recall rate, representing 22% of the 16-second advertisement). Table 8 compares brand disclosure timing across advertisements and its relationship to engagement outcomes, revealing the complex relationship between commercial transparency and engagement sustainability.

Table 8 *Brand Disclosure Timing and Engagement Outcomes*

Advertisement	Brand Disclosure Timing	Time to Brand Recognition (s)	% Participants Recognizing Early	Mean Viewing Time (s)	% Viewing Beyond Recognition	Interview Sentiment
Oatly Mermaid (post 77)	Immediate (0-2s)	1.8	89%	3.59	74%	Positive creative appreciation

Zalando (post 104)	Immediate (0-2s)	1.2	92%	1.13	18%	Negative manipulation perception
Oatly Lab (post 82)	Delayed (8s+)	8.3	15% (before 8s)	6.93	83%	Positive respect for creativity
Travel (post 143)	Intermediate (3-6s)	Variable (5-12s)	34%	5.70	91%	Positive authenticity perception
Lipton (post 41)	Immediate (0-1s)	0.8	95%	1.57	12%	Negative aggressive perception

Note. s = seconds. Time to Brand Recognition measured from advertisement start to point when brand identity became apparent to majority of participants (based on interview recall). % Participants Recognizing Early indicates proportion of participants who viewed the advertisement who identified commercial nature within first 3 seconds (N=19 for Travel and Oatly Mermaid; N=20 for Oatly Lab, Lipton, and Zalando). % Viewing Beyond Recognition shows proportion who continued viewing 2+ seconds after brand recognition occurred. Interview Sentiment summarizes predominant evaluative tone in participant responses.

Eye-tracking data showed that early brand visibility in Oatly Mermaid did not significantly reduce fixation duration or trigger immediate scrolling when paired with visually novel creative execution. Fixation durations remained stable at 298ms average throughout viewing—comparable to the curiosity-driven Oatly Lab (287ms) and only marginally lower than the emotionally coherent Travel (312ms), and dramatically higher than the early-brand-disclosure Lipton (134ms) and Zalando (156ms). The critical difference

appeared to be not brand disclosure timing per se, but the presence of creative value that justified continued attention despite commercial recognition.

Gaze pattern analysis revealed focused rather than scattered viewing in Oatly Mermaid, characterized by concentrated attention on the central mermaid figure (68% of total viewing time), systematic visual exploration of unusual costume elements and beach setting (22% of viewing time), and minimal avoidance behaviors with only 2.1 saccadic movements per second compared to 4.2-4.7 in rejected advertisements. These patterns suggested that creative novelty sustained attention even when commercial intent was immediately apparent, creating engagement through aesthetic interest and creative appreciation rather than through narrative investment or emotional transportation.

Participants responded positively to the combination of commercial transparency and creative risk-taking, describing appreciation for honest advertising that respected their intelligence while providing entertainment value through unexpected creative execution. "They weren't hiding it, but it was still interesting. Like who thinks of mermaids for milk advertising?" (P09) observed one participant, explicitly valuing both the transparent commercial framing and the creative unexpectedness—suggesting that creativity can justify commercial content when executed with sufficient originality. Another explained the creative justification: "It was so random but lowkey worked. At least they were creative" (P05), using language suggesting that creative originality compensated for commercial intent, creating an implicit exchange where creative value justified attention investment. A third elaborated on the honesty appreciation: "I couldn't even tell if it was an ad or like where it was going. It was new" (P13), describing how creative novelty created

its own form of curiosity distinct from narrative-driven approaches—a "what will they do next" interest rather than "what happens next" narrative curiosity.

These responses indicate that authenticity judgments were based on creative originality and departure from conventional advertising expectations rather than on concealment of commercial intent, suggesting a more sophisticated evaluation process than simple resistance to recognized advertising. This suggests that Generation Z's advertising literacy extends beyond mere recognition of commercial content to evaluative judgments about the honesty and quality of commercial approaches.

Several participants explicitly appreciated the honest commercial framing paired with unexpected creative choices, framing this combination as a form of respect for audience intelligence that enhanced rather than diminished receptivity. "Gotta give them credit for trying something creative instead of boring" (P20) exemplified responses that valued creative risk-taking as a positive brand signal indicating confidence and audience respect. Another elaborated: "It wasn't like other ads, so that's pretty cool. They tried something different" (P02), describing creative departure from conventional approaches as inherently valuable regardless of product relevance or personal interest. A third made the respect assessment explicit: "When a brand is weird and creative like that, it shows they don't think we're idiots who will just buy anything. They're trying to entertain us, not just sell to us" (P17). These responses suggest that strategic authenticity—honest commercial framing paired with creative value—may resonate particularly well with Generation Z audiences who have developed sophisticated advertising literacy through extensive social media exposure.

Eye-tracking patterns confirmed that creative novelty facilitated processing despite commercial recognition. Focused rather than scattered viewing occurred in 78% of viewing time, with sustained fixations on the central mermaid figure rather than rapid scanning suggesting comfortable engagement. Systematic visual exploration of unusual elements showed viewers methodically examining costume details, beach setting, and product placement in coordinated sequences indicating active appreciation rather than passive viewing. Minimal avoidance behaviors were evident, with low saccadic frequency (2.1 per second) and no fixations on interface elements or surrounding content—a pattern associated with exit-seeking in rejected advertisements. This processing pattern suggests that creative novelty sustained attention through aesthetic interest even when commercial intent was immediately apparent, creating conditions where viewers willingly engaged with advertising content because it provided entertainment or creative value above typical commercial messaging.

The lower standard deviation for Oatly Mermaid (SD = 2.96 seconds) compared to Oatly Lab (SD = 5.85 seconds) indicated more consistent engagement across participants, suggesting that strategic authenticity may produce more predictable outcomes than curiosity-driven narratives though with somewhat shorter absolute viewing times. While curiosity-driven approaches created bimodal distributions with some viewers deeply engaged and others quickly disengaging, strategic authenticity generated more consistent moderate engagement across the participant pool. Analysis showed 74% of participants (14 of 19 who viewed the advertisement) crossed the three-second threshold, compared to 65% for Oatly Lab, though fewer sustained viewing beyond 8 seconds (26% versus 45% for Oatly Lab). This pattern suggests strategic authenticity may provide more reliable baseline engagement while curiosity-driven approaches offer higher ceiling potential with greater risk.

The success of strategic authenticity challenges traditional advertising approaches emphasizing subtle integration and delayed brand disclosure, suggesting that transparency about commercial intent does not inherently impede engagement when paired with sufficient creative value. This finding has important implications for understanding what drives positive versus negative responses, as explored in RQ3, where the relationship between commercial transparency and viewer resistance is examined more directly through comparison of successful transparent approaches (Oatly Mermaid) with failed concealment attempts (Zalando) and overly obvious approaches lacking creative value (Lipton).

Emotional Coherence Through Coordinated Design Elements

Beyond curiosity-driven narratives and strategic authenticity, a third distinct mechanism emerged for maintaining engagement: emotional coherence through coordinated design elements. The Travel advertisement (post 143) achieved sustained engagement through systematic coordination of visual, auditory, and editorial elements that created immersive emotional experiences. This represents a third distinct mechanism: holistic design integration rather than narrative development or creative novelty. With a 5.70-second average viewing time representing 41% of the 14-second advertisement duration—the highest completion percentage among analyzed advertisements—this content demonstrated how emotional coherence maintains attention more effectively than either curiosity-driven narratives or creative novelty alone, achieving both high absolute engagement and high proportional completion. The

advertisement's success stemmed not from any single exceptional element but from the systematic coordination of multiple components that worked together to create a unified emotional experience characterized by processing ease and immersive absorption.

Table 9 presents detailed analysis of design element coordination in Travel compared to uncoordinated approaches, revealing how systematic integration creates qualitatively different viewing experiences.

Table 9 *Design Element Coordination and Engagement Quality*

Element Category	Travel (Coordinated)	Lipton (Uncoordinated)	Coordination Indicators	Impact on Engagement
Visual aesthetics	Soft natural palettes, dynamic scenes, smooth transitions	Saturated colors, rapid changes	Color harmony, transition smoothness, scene coherence	Comfortable vs. overwhelming
Audio elements	Moderate tempo, balanced volume, looping pattern	High tempo, loud volume, abrupt start	Tempo matching, volume appropriateness, musical flow	Pleasant vs. aggressive
Editorial rhythm	Cuts synchronized with beats, consistent pacing, smooth scene flow	Rapid pacing	Beat synchronization, pacing consistency, scene connection	Immersive vs. disjointed

Element integration	Visual, audio, editorial working together	Elements competing for attention	Cross-modal coordination	Coherent vs. chaotic
Viewing time	5.70s (41% completion)	1.57s (26% completion)	Sustained vs. Abandoned	High vs. low engagement
Fixation pattern	Stable (312ms), coordinated	Scattered (134ms), irregular	Processing quality	Deep vs. superficial
Emotional response	"Soft break in feed," pleasant absorption	"Too aggressive," aversive arousal	Affective quality	Positive vs. negative

Note. All metrics are based on participants who viewed each advertisement (N=19 for Travel; N=20 for Lipton). Coordination Indicators describe key metrics of element integration. Impact categories summarize effects on viewer experience.

Eye-tracking analysis revealed engagement indicators distinct from both curiosity-driven and novelty-based strategies, suggesting a qualitatively different form of attention characterized by immersion rather than active information-seeking. Patterns included reduced gaze dispersion with 82% of fixations concentrated in central content area rather than scattered across advertisement space, indicating focused attention on primary content rather than exploratory scanning. Stable viewing patterns showed consistent 312ms fixation durations throughout viewing with only 2% variation (309ms-318ms range across different viewing phases), suggesting comfort and immersion rather than effortful processing or declining interest. Coordinated eye movements synchronized with editorial rhythm demonstrated viewers' gaze following visual flow and transitioning between scenes in coordination with cuts and

audio beats, indicating absorbed viewing rather than independent visual scanning. Minimal saccadic frequency of only 1.8 rapid eye movements per second—the lowest among all advertisements—suggested visual comfort and absence of search behavior for alternative content or exit points.

These patterns suggest that viewers were experiencing the content holistically rather than processing individual elements sequentially, creating what might be characterized as absorbed or immersed viewing similar to flow states described in media engagement research. Gaze plots revealed smooth, flowing eye movement patterns that followed natural visual composition and editorial direction rather than random or systematic scanning, suggesting viewers were allowing the advertisement to guide their attention rather than actively directing it—a characteristic of immersive rather than critical engagement.

Participants described the viewing experience in notably emotional terms that emphasized feeling over thinking, using affective language rather than cognitive descriptors. "It felt like old times—even if it wasn't really something I've done myself" (P01) described one participant, using nostalgic language suggesting emotional transportation to imagined experiences rather than rational evaluation. Another used absorption language: "Made me wanna live those moments. It was like a soft break in the feed" (P09), describing the advertisement as creating desirable emotional states and providing respite from typical content consumption—a characterization suggesting the content functioned more like entertainment than advertising. A third emphasized the immersive quality: "I got lost in it for a second. It was like watching a friend's travel video, not an ad" (P06), explicitly noting diminished critical awareness characteristic of immersive experiences where advertising recognition fades into background.

These emotional descriptions contrast sharply with the cognitive language used for curiosity-driven content ("trying to figure it out," "wanted to understand") and creative appreciation language for strategic authenticity ("that's creative," "respect the originality"), suggesting that emotional coherence operates through affective engagement rather than cognitive or aesthetic mechanisms. Participants weren't actively thinking or evaluating—they were feeling and experiencing, a distinction with important implications for understanding engagement quality beyond mere duration.

The advertisement's success derived from coordinated design strategy rather than isolated emotional appeals, with multiple elements working together systematically to create coherent experience. Visual components included soft natural color palettes (blues, greens, warm earth tones) creating calm emotional tone rather than high-intensity stimulation that might trigger arousal or overwhelm. Dynamic activity scenes (people swimming, jumping, exploring) maintained visual interest and movement without creating chaos or requiring intense cognitive processing. Smooth camera transitions using fades and pans reduced cognitive disruption that might break immersion or require reorientation. Bright natural lighting enhanced positive affect through associations with outdoor recreation and pleasant environments. These visual choices worked together to create an aesthetically pleasing, emotionally positive viewing environment that felt natural rather than constructed.

The original audio features were designed to foster emotional coherence and processing ease. This was achieved through several strategic elements. Audio elements featured looping background music with a positive emotional valence (major key, uplifting melody), creating a consistent emotional tone throughout the advertisement. Crucially, the moderate tempo matched the visual rhythm

without generating a sense of rush or urgency that might induce stress. The use of familiar audio—popular TikTok sounds or trending music—created recognition and comfort through association with positive previous exposures. Finally, balanced volume levels prevented sensory overload while maintaining presence and emotional impact. This precise audio coordination with visual elements created audiovisual synchrony—the perception that sound and image work together naturally—which research associates with processing ease and positive affect.

Equally critical, editorial rhythm was carefully managed through two main techniques. First, it incorporated rapid cuts synchronized with musical beats (cuts occurring on or just before beat emphasis), which created momentum and maintained attention through rhythmic coordination. Second, smooth transitions between visual scenes—using fades or motion-matched cuts—prevented jarring interruptions that might break immersion or require cognitive reorientation. It was this precise editorial coordination that participants experienced as flow—the perception that the advertisement progressed naturally and inevitably rather than through constructed editing.

Participants explicitly valued this coordinated approach, describing how the integration of elements created qualitatively superior experience compared to advertisements with strong individual elements but poor coordination. "Everything just flowed together—the music, the scenes, the vibe. It felt good" (P04) exemplified responses emphasizing how elements worked together rather than individual components, with participants describing the advertisement as feeling 'complete' or 'cohesive Another made the

coordination assessment explicit: "It wasn't trying too hard. Everything matched and felt natural" (P19), contrasting coordinated integration with forced or obvious design decisions that call attention to themselves. A third described the processing ease: "I didn't have to think about it—it just felt right. Like I could relax and enjoy it instead of analyzing it" (P14), explicitly noting the reduced cognitive load and enhanced affective experience created by good coordination.

Several participants expressed behavioral intentions resulting from this emotional engagement, indicating that emotional coherence translated into concrete outcomes beyond mere viewing completion. "I want to send this to my friend so she gets FOMO and travels with me" (P17) demonstrated sharing intention motivated by emotional resonance and desire to create shared experience. "It reminded me of what I want to do—it kind of hit a nerve in a good way" (P10) described how emotional connection translated into personal aspiration and potential behavioral motivation. "I could see myself saving this for travel inspiration" (P04) indicated perceived utility value extending beyond commercial messaging to genuine content value worthy of preservation. A fourth participant described consideration development: "I don't know what it was advertising exactly, but whatever it is, I'm interested now because of how it made me feel" (P08), explicitly connecting emotional response to commercial outcomes despite unclear brand messaging—suggesting that emotional engagement can create brand interest even absent clear product communication.

These behavioral intentions distinguish emotional coherence from other strategies in important ways. While curiosity-driven approaches generated appreciation and positive brand associations ("I respect that they got me interested"), and strategic authenticity created creative appreciation ("at least they were creative"), emotional coherence appeared to create stronger motivation for concrete

action including content sharing, behavioral change consideration, and active information-seeking about the advertised offering. This suggests that emotional coherence may be the most effective strategy for generating behavioral outcomes beyond passive viewing and positive attitude formation, though the current study's focus on immediate responses rather than actual behavioral tracking limits confidence in this conclusion.

The relatively low standard deviation (SD = 3.88 seconds) compared to curiosity-driven approaches (SD = 5.85) indicated consistent engagement across participants, suggesting that emotional coherence produces more predictable outcomes than strategies relying on individual differences in curiosity or appreciation for creative novelty. Analysis showed 91% of participants (17 of 19 who viewed the advertisement) crossed the three-second threshold—the highest proportion among all advertisements—and 68% continued beyond 5 seconds, demonstrating broad appeal with minimal polarization. This consistency, combined with the high completion percentage (41%) and strong behavioral intentions, positioned emotional coherence as perhaps the most reliable engagement maintenance strategy among those analyzed, achieving both high engagement levels and broad appeal with minimal risk of complete failure.

However, understanding why emotional coherence succeeded requires examining what happened when coordination failed and design elements competed rather than cooperated—a pattern exemplified by the Lipton advertisement's high-intensity sensory approach that achieved rapid attention capture but immediate rejection.

Contrast: High-Intensity Sensory Approaches and Engagement Failure

Lipton's (post 41) high-intensity sensory strategy demonstrated that sensory impact does not equate to engagement maintenance, providing critical contrast to the emotional coherence approach and revealing how uncoordinated intensive stimulation can trigger defensive responses rather than engagement. Despite achieving rapid initial attention capture with an 89% first-fixation rate on the central brand logo within 290ms—the most efficient attention capture among all analyzed advertisements—the intensive approach triggered immediate negative responses and rapid disengagement. With only 1.57 seconds average viewing time—less than one-quarter of the 6-second advertisement duration—Lipton demonstrated the shortest sustained engagement among analyzed advertisements, creating a stark inverse relationship between attention capture efficiency and engagement maintenance success. Table 10 presents a systematic comparison of sensory intensity characteristics and their relationship to engagement outcomes.

Table 10 *Sensory Intensity Comparison: Coordinated vs. Uncoordinated Approaches*

Characteristic	Travel (Coordinated)	Lipton (Uncoordinated High-Intensity)	Effect on Engagement
Visual intensity	Moderate, varied	High, constant saturation	Comfortable vs. overwhelming
Color saturation	Natural tones (40-60% saturation)	Highly saturated (80-95% saturation)	Pleasant vs. harsh

Audio volume	Moderate	High, abrupt start	Comfortable vs. Jarring
Audio tempo	Moderate	High	Relaxed vs. Rushed
Visual changes	Quick (1.0s avg scene)	Rapid	Processable vs. Overwhelming
Element coordination	Synchronized	Messy	Coherent vs. Chaotic
Mean viewing time	5.70s	1.57s	Sustained vs. Abandoned
Fixation duration	312ms	134ms	Deep vs. Superficial
Saccadic frequency	1.8 per second	4.7 per second	Comfortable vs. Stressed
Emotional response	Pleasant absorption	Viewer Irritation	Positive vs. Negative
Behavioral intention	Sharing, interest	Avoidance, escape	Approach vs. avoid

Note. s = seconds; ms = milliseconds. All metrics based on participants who viewed each advertisement (N=19 for Travel; N=20 for Lipton). Saturation percentages estimated from color analysis. Audio volume jump measured relative to preceding content. Avg = average duration.

Participants reported strong negative reactions to the sensory intensity using language indicating aversive arousal rather than positive stimulation—describing experiences more similar to assault than appeal. "It came on too strong—I was out of it in seconds" (P14) exemplified immediate rejection responses using intensity language suggesting overwhelm rather than mere disinterest. Another used aggressive language: "The sound was annoying, like aggressive. I just wanted it to stop" (P06), explicitly framing sensory intensity as violating personal space or comfort rather than merely being unappealing—a characterization suggesting defensive

responses. A third described physical discomfort: "Too much happening at once. It hurt my eyes" (P18), using pain language indicating that sensory intensity crossed from stimulation to actual physical discomfort, creating aversive experience triggering avoidance.

These descriptions suggest that high-intensity sensory approaches may trigger defensive responses similar to fight-or-flight reactions to overwhelming stimulation rather than the positive arousal that advertisers presumably intended to create. The language choice—"aggressive," "hurt," "assault," "escape"—indicates that participants experienced Lipton as threat rather than appeal, creating conditions where immediate disengagement functioned as self-protection rather than preference expression.

Eye-tracking data confirmed reported discomfort through multiple behavioral indicators distinct from normal disengagement patterns. Shortened fixation durations averaging only 134ms—less than half the duration of successful advertisements (287-312ms)—indicated reduced willingness to process visual information and minimal cognitive investment in content comprehension. Increased saccadic frequency of 4.7 rapid eye movements per second—more than double successful advertisements (1.8-2.3 per second) and the highest among all analyzed content—suggested visual stress and active scanning for escape routes rather than content exploration. Irregular scan patterns characterized by scattered, non-systematic eye movements covering advertisement periphery and interface elements (particularly scroll area) indicated active avoidance rather than engaged processing. Gaze plots revealed attention concentrated on exit affordances (scroll bar, surrounding content) rather than central advertisement content, a pattern researchers associate with escape-seeking behavior.

The contrast between Lipton's sensory intensity and the Travel advertisement's emotional coherence was particularly instructive for understanding what distinguishes successful from unsuccessful sensory approaches, as both employed bright visuals, dynamic editing, and upbeat audio—seemingly similar sensory elements on surface examination. However, their coordination and intensity differed dramatically in ways that produced opposite engagement outcomes. Travel's coordinated approach featured moderate intensity with systematic synchronization across sensory channels, creating immersive emotional experience that viewers described as pleasant and absorbing—"like a soft break in the feed" that provided respite from typical content consumption patterns. Visual elements remained processable with scene durations allowing comprehension, color palettes remained comfortable with natural saturation levels (40-60%), and audio maintained moderate volume with gradual transitions preventing jarring shifts. The coordination created processing ease where viewers could comfortably absorb information without cognitive strain.

In stark contrast, Lipton's uncoordinated intensity featured high-volume audio that started abruptly without gradual entry, creating a jarring jump from preceding content that participants described as "aggressive" and "attacking." Most critically, uncoordinated elements that competed for attention rather than working together—audio tempo mismatched to visual transition rate, color choices clashing rather than complementing, editing rhythm disconnected from audio beats—created chaotic rather than immersive experience requiring active management of competing stimuli. This contrast suggests that effective engagement maintenance requires systematic coordination of design elements at appropriate intensity levels rather than maximization of sensory impact across all channels simultaneously. The distinction between "coordinated moderate intensity" and "uncoordinated high

intensity" appears more important than absolute intensity levels, as moderate elements working together create stronger engagement than intense elements competing against each other.

The failure of high-intensity approaches despite successful attention capture reveals an important principle: capturing attention and maintaining engagement require fundamentally different strategies, with approaches optimized for one often working against the other. Attention capture favors high contrast, sudden changes, and intense stimuli that break through perceptual filtering and force attention allocation. Engagement maintenance favors processing ease, comfortable stimulation levels, and coordinated experiences that enable sustained attention without cognitive strain or discomfort. Advertisements maximizing attention capture through sensory intensity may simultaneously minimize engagement maintenance by creating processing difficulty and autonomy threats that trigger rapid disengagement once attention is captured.

This principle connects directly to RQ3's examination of what determines positive versus negative responses, as the factors driving rapid disengagement from sensory-intense advertisements illuminate broader patterns of advertising resistance. The sensory overwhelm, autonomy violation, and processing difficulty that characterized negative responses to Lipton represent specific instances of broader reactance triggers that determine whether engagement translates into positive outcomes or defensive rejection—patterns explored systematically in the following section.

RQ3: Factors Determining Positive Versus Negative Responses

Having identified what captures initial attention (RQ1) and what maintains engagement beyond the three-second threshold (RQ2), the third research question examines the crucial final stage: what determines whether users respond positively or negatively to branded content. This question explores why some advertisements that successfully captured attention—like Lipton and Zalando—ultimately triggered avoidance, while others—like the Oatly advertisements and Travel—generated favorable brand associations and behavioral intentions. Analysis revealed that response valence was determined primarily by three interrelated factors: the critical three-second engagement threshold distinguishing commitment from abandonment, processing fluency reflecting ease of interpretation within platform conventions, and the presence or absence of psychological reactance triggers that activated defensive responses.

The Critical Three-Second Engagement Threshold

The three-second engagement threshold identified in the Results extends Lang's (2000) Limited Capacity Model, functioning as a commitment decision point with profound implications for ultimate response valence. Advertisements maintaining attention beyond three seconds demonstrated substantially higher completion rates, generated positive recall associations, and produced favorable brand evaluations, while advertisements abandoned within three seconds generated minimal recall or predominantly negative associations. This threshold appeared to represent a transition point where initial visual interest must successfully convert to deeper cognitive or emotional investment, separating content that merely captures attention from content that earns continued engagement.

Eye-tracking data revealed distinct patterns differentiating threshold crossing from early abandonment, providing behavioral evidence for this cognitive decision point. For advertisements exceeding three seconds (Oatly Lab: 6.93s, Travel: 5.70s, Oatly Mermaid: 3.59s), four patterns indicated sustained engagement. First**,** fixation durations averaged 287-312ms, indicating systematic processing rather than superficial scanning. Second**,** gaze patterns remained focused rather than scattered, suggesting engaged attention directed toward meaning-making. Third**,** viewers demonstrated systematic visual exploration, indicating active information-seeking or aesthetic appreciation. Fourth**,** attention showed minimal decline following initial engagement, suggesting that the decision to continue viewing was stable rather than continuously re-evaluated. These patterns indicate qualitatively different cognitive processing compared to brief exposures. Crossing the three-second threshold appears to involve transitioning from evaluation mode (deciding whether content merits attention) to engagement mode (processing content for meaning or enjoyment).

Participants described crossing this threshold as a commitment decision marked by reduced critical evaluation and increased open reception. "Once I'm a few seconds in and it's still interesting, I'll usually finish it or at least watch more" (P08) explained one participant, explicitly describing the threshold as a decision point that reduced subsequent likelihood of abandonment. Another noted, "If something grabs me in the first couple seconds and keeps me there, I'm probably staying" (P05), using language suggesting that early engagement success creates momentum toward completion. These descriptions suggest that the three-second threshold

represents not just a temporal marker but a psychological transition from tentative sampling to committed viewing, with important implications for how advertising content should be structured.

Conversely, advertisements abandoned within three seconds (Lipton: 1.57s, Zalando: 1.13s) exhibited different eye-tracking patterns consistent with rejected rather than merely incomplete viewing. These patterns included shortened fixation durations (134-156ms) indicating reduced processing depth and unwillingness to invest cognitive resources, increased saccadic movement suggesting visual discomfort or active scanning for escape, irregular scanning patterns consistent with avoidance behavior rather than systematic exploration, and rapid scrolling initiation often beginning before the three-second mark was reached. These behavioral indicators suggest active rejection rather than passive abandonment, indicating that failure to cross the threshold involved not merely insufficient appeal but active negative response.

Participants described early abandonment decisions as immediate rejection based on rapid negative assessment, using language suggesting active dismissal rather than neutral disinterest. "I could tell in like two seconds it wasn't worth my time" (P03) explained one participant, framing the decision as value judgment rather than preference. Another stated, "As soon as I saw it was just trying to sell me something, I scrolled" (P11), explicitly linking commercial recognition to rejection behavior. These descriptions suggest that failing to cross the three-second threshold often involves triggering negative associations or resistance rather than simply failing to generate sufficient interest, a distinction with important implications for understanding response valence.

The three-second threshold appeared to represent a decision point where initial visual appeal must successfully transition to sustained cognitive investment through one of the mechanisms identified in RQ2: curiosity generation (Oatly Lab), emotional resonance (Travel), or creative interest (Oatly Mermaid). Advertisements relying solely on visual salience without deeper engagement mechanisms (Lipton, Zalando) failed to sustain attention beyond initial exposure, never allowing the transition from evaluation to engagement mode. This finding connects the three stages of engagement—initial attention capture, engagement maintenance, and response outcomes—revealing that success at each stage depends on successfully transitioning to the next rather than simply maximizing performance at any single stage. Understanding what facilitates or impedes this transition requires examining the role of processing fluency in creating smooth versus disrupted viewing experiences.

Processing Fluency and Platform Integration

Positive responses were consistently associated with high processing fluency—the subjective ease with which participants could interpret and connect with content within TikTok's platform conventions—while negative responses corresponded with processing difficulty and perceived platform misalignment. The Travel advertisement exemplified successful platform integration, achieving the highest viewing completion percentage (41%) through visual aesthetics, audio selection, and editorial rhythms indistinguishable from organic user-generated content. This seamless integration created what participants experienced as effortless viewing, where content interpretation required minimal cognitive resources and felt natural within the platform context.

Participants explicitly valued this seamless platform integration, describing how it shaped their receptivity to commercial messaging by reducing rather than triggering commercial recognition. "It didn't scream 'ad'—it just fit with everything else I was watching" (P07) explained one participant, suggesting that platform alignment delayed or prevented the commercial recognition that might activate resistance. Another noted, "I honestly thought it was someone's actual travel video at first" (P15), indicating that successful integration creates ambiguity about content type that works in advertisers' favor by allowing open reception rather than defensive processing. Eye-tracking data supported these subjective reports through behavioral indicators of processing ease: smooth, coordinated gaze patterns rather than scattered scanning suggesting comfortable navigation of visual information, longer sustained fixations indicating willingness to invest attention in content processing, and reduced saccadic movement suggesting visual comfort and absence of search behavior for alternative content.

The concept of processing fluency extended beyond mere visual aesthetics to encompass what participants described as cultural literacy—demonstration of understanding platform-specific conventions and user expectations that goes beyond surface-level aesthetic mimicry. Participants distinguished between advertisements that merely looked like TikTok content and those that demonstrated genuine understanding of platform culture, with the latter generating more positive responses. "They actually understand TikTok, not just using it to advertise" (P12) explained one participant, drawing a distinction between authentic platform fluency and superficial imitation. Another elaborated, "It feels like they're part of the TikTok page and vibe, not just interrupting it" (P20), using language suggesting that successful platform integration requires demonstrating insider status rather than merely adopting external markers. This cultural dimension of processing fluency suggests that effective platform integration requires deeper engagement with

platform norms than simple aesthetic matching, involving understanding of communication styles, content formats, and community values.

This finding has important implications for understanding why some advertisements succeeded while others failed: processing fluency appears to operate as a mediating factor that translates platform integration into positive response by reducing cognitive friction and resistance. Advertisements achieving high processing fluency created conditions where commercial intent could coexist with positive response, while those creating processing difficulty triggered negative associations regardless of other qualities they might possess. This connection between processing ease and positive/negative responses helps explain why factors identified in RQ1 and RQ2—such as platform-native aesthetics and emotional coherence—influenced ultimate outcomes: they succeeded partly by facilitating processing fluency that enabled rather than impeded engagement. Understanding the opposite pattern—how processing difficulty contributed to negative responses—requires examining the specific triggers that activated defensive rather than receptive processing.

Psychological Reactance Triggers and Negative Responses

Negative responses were consistently triggered by specific elements that participants perceived as violating platform norms, manipulating their attention, or disrespecting their intelligence, activating psychological reactance that transformed potential engagement into active resistance. The two negatively recalled advertisements (Lipton and Zalan do) both featured multiple reactance

triggers that generated defensive responses, providing clear contrast to the processing fluency and platform integration that characterized successful advertisements. Analysis revealed four primary reactance triggers that consistently produced negative responses: overly obvious commercial intent, aesthetic misalignment with platform conventions, sensory overwhelm through aggressive approaches, and perceived manipulation through deceptive authenticity attempts.

Overly Obvious Commercial Intent

Prominent brand logos in opening frames triggered immediate commercial recognition that activated critical evaluation rather than open reception, fundamentally altering viewers' interpretive frame from the first moments of exposure. "The big Lipton logo right in your face—they're not even trying to be subtle" (P14) explained one participant, framing prominent branding as a respect violation rather than merely a commercial disclosure. Another elaborated, "When the brand is that obvious from the start, I know I'm being sold to and I don't want to engage" (P06), explicitly describing how early brand prominence activated defensive processing that precluded receptive viewing. Eye-tracking data confirmed that while prominent logos captured initial attention rapidly (290-310ms first fixation)—the fastest attention capture among all analyzed advertisements—they also corresponded with the shortest sustained viewing times (1.13-1.57 seconds) and most rapid scrolling initiation, creating an inverse relationship between branding efficiency and engagement success.

The timing of brand disclosure emerged as particularly critical, with early prominence triggering resistance while delayed or integrated disclosure allowed engagement to develop before commercial recognition occurred. Comparing Oatly Lab (brand revealed at 8 seconds, 6.93s average viewing) with Lipton (brand visible immediately, 1.57s average viewing) illustrates this timing effect: identical brand obviousness produced opposite outcomes depending on when commercial intent became apparent. This suggests that the issue is not brand visibility per se but rather whether viewers have opportunity to develop interest or investment before commercial recognition activates critical evaluation. Early brand prominence prevents this investment from forming, creating conditions where viewers encounter commercial intent before experiencing any value that might justify continued attention.

Aesthetic Misalignment with Platform Conventions

High-production values that contrasted sharply with typical user-generated content immediately signaled commercial intent, even before brand elements appeared. The polished, professional aesthetic made advertisements stand out as 'not TikTok,' revealing their commercial nature through visual mismatch alone. "It looked like a TV commercial, not a TikTok. Wrong place, wrong vibe" (P18) explained one participant, explicitly contrasting the advertisement's production quality with platform expectations. Another elaborated, "Everything was too polished, too perfect. It didn't fit" (P11), describing how aesthetic deviation from platform norms created dissonance that highlighted commercial nature rather than masking it. Participants described this misalignment as demonstrating lack of platform understanding and respect for community norms: "They just took a regular ad and put it on TikTok without thinking about what works here" (P09), framing aesthetic mismatch as evidence of outsider status that justified rejection.

This aesthetic mismatch created a double problem. First**,** the professional look immediately revealed commercial intent. Second**,** it showed the brand didn't understand TikTok's culture. Participants felt both manipulated—brands were trying to disguise ads—and disrespected—brands failed to understand the platform. "It's like they think we're too dumb to notice it doesn't belong here" (P16) explained one participant, describing the mismatch as insulting rather than just unappealing. These strong negative reactions help explain why informal aesthetics succeeded: they avoided triggering resistance by demonstrating cultural understanding and respecting platform conventions.

Sensory Overwhelm and Aggressive Approach

High-intensity sensory elements—particularly sudden loud audio and highly saturated visual elements—triggered active avoidance rather than engagement, creating aversive arousal that participants described using language suggesting physical discomfort or assault. "The sound was so aggressive it made me want to get away from it" (P06) explained one participant, framing sensory intensity as violation of personal space rather than merely excessive stimulation. Another elaborated, "Too bright, too loud, too much—I had to scroll just to escape" (P14), explicitly describing scrolling as escape behavior rather than neutral content rejection. Eye-tracking data confirmed avoidance behavior through increased saccadic frequency indicating visual stress and scanning for escape routes, shortened fixation durations suggesting unwillingness to maintain visual contact with uncomfortable stimuli, and irregular scanning patterns consistent with defensive rather than exploratory viewing.

The aggressive nature of sensory intensity appeared particularly problematic in TikTok's context, where users exercise complete control over content exposure and expect smooth, user-controlled experiences. Sudden loud audio or visually overwhelming content violated expectations of user control, creating perception of forced attention that triggered reactance responses designed to reassert autonomy. "I'm choosing what to watch—when something forces itself on me like that, I'm immediately out" (P13) explained one participant, explicitly framing sensory intensity as autonomy threat. This finding helps explain why emotional coherence succeeded in RQ2 while sensory intensity failed: coordinated moderate-intensity design respected user autonomy and control expectations, while high-intensity approaches violated these expectations and triggered defensive responses aimed at reasserting control through rapid scrolling.

Perceived Manipulation Through Deceptive Authenticity

Several participants noted feeling manipulated by advertising techniques they recognized as intentionally deceptive or exploitative, particularly when advertisements attempted to appear authentic while employing obvious commercial strategies. Regarding Zalando's attempt at relatability through young creators and casual setting combined with obvious commercial elements, one participant observed: "They're trying to trick you into thinking it's authentic, but you can tell it's fake" (P16), describing the attempt as patronizing manipulation. Another noted, "It's insulting when they think you won't notice it's an ad just because they hired someone Gen-Z" (P13), framing the strategy as intelligence insult that activated resistance stronger than honest commercial framing would have triggered.

This perceived manipulation triggered particularly strong resistance because it combined commercial intent (which participants could tolerate) with perceived deception (which they found offensive), creating a double violation worse than obvious advertising. Comparing responses to Oatly Mermaid (obvious commercial intent, creative honesty, 3.59s viewing) with Zalando (attempted authenticity with obvious commercial elements, 1.13s viewing) illustrates how transparent commercial framing generated more positive responses than attempted concealment recognized as deceptive. Participants repeatedly expressed preference for honest advertising that respected their intelligence over manipulative approaches that assumed they could be easily fooled: "Just be honest that it's an ad and make it interesting—don't pretend to be something you're not" (P11) exemplified this preference for transparent commercial intent over deceptive authenticity attempts, a sentiment echoed by multiple participants (P08, P16).

This finding challenges traditional advertising wisdom emphasizing subtle integration and delayed disclosure, suggesting that in contexts where users have developed sophisticated advertising literacy—as Generation Z has through extensive social media exposure—transparency may generate more positive responses than attempted concealment. The key appears to be pairing transparency with actual value—creativity, entertainment, or emotional resonance—rather than relying on concealment alone to prevent resistance. Advertisements that achieved this balance (Oatly Mermaid, to some extent Oatly Lab) generated positive responses despite obvious commercial intent, while those attempting concealment that failed (Zalando) triggered stronger resistance than honest commercial framing would have produced.

Positive Behavioral Intentions and Brand Interest

Advertisements that successfully navigated the three-second threshold, achieved processing fluency, and avoided reactance triggers generated positive behavioral intentions that extended beyond mere viewing completion. These positive responses took several forms, ranging from brand appreciation and positive associations to sharing intentions and purchase consideration. For Oatly Lab, participants expressed appreciation for creative risk-taking that translated into positive brand associations, even without immediate purchase intent. "I don't even drink oat milk, but I respect the creativity. Makes me think better of the brand" (P02) explained one participant, indicating that creative execution created brand value independent of product relevance. Another elaborated, "If they're willing to be that weird and creative, they're probably cool as a company" (P20), describing how advertising creativity served as brand character signal that influenced overall brand perception. These responses suggest that effective engagement builds brand equity through positive associations, even without immediate conversion.

The Travel advertisement generated the strongest expressions of behavioral intent among all analyzed advertisements, with multiple participants indicating concrete actions they would or might take as a result of viewing. "I want to send this to my friend so she gets FOMO and travels with me" (P17) exemplified sharing intentions that would extend advertisement reach beyond initial exposure. "I'd probably click through to see what app or service this is" (P10) indicated active interest in obtaining more information about the advertised offering. "I could see myself saving this for travel inspiration" (P04) described how the content provided value beyond commercial messaging that justified preservation for future reference. These diverse behavioral intentions—sharing,

information-seeking, content preservation—demonstrate that effective emotional engagement translates into concrete outcomes beyond passive viewing.

Notably, positive responses did not always translate into immediate purchase intention but rather into favorable brand perception and willingness to engage further, suggesting that short-form advertising effectiveness should be evaluated across multiple outcome dimensions rather than conversion alone. "I'm not booking a trip right now, but I'm definitely more interested in whatever they're promoting" (P08) explained one participant, describing increased consideration rather than immediate action. Another noted, "It made me curious enough that I'd look them up later" (P19), indicating that engagement created openness to future interaction rather than requiring immediate response. These graduated positive outcomes suggest that effective short-form advertising creates multiple pathways to value—immediate conversion for high-intent users, increased consideration for moderate-intent users, and positive brand associations for low-intent users—rather than requiring single-outcome success.

The pattern of positive responses reveals an important principle: successful advertisements created perceived value that justified attention investment, whether through entertainment (Oatly Lab), emotional resonance (Travel), or creative novelty (Oatly Mermaid). This value provision appeared to function as implicit exchange that legitimized commercial messaging—participants tolerated or even welcomed advertising when they received something valuable in return. Conversely, advertisements failing to provide value beyond commercial messaging (Lipton, Zalando) were perceived as one-sided exchanges that demanded attention without offering adequate return, triggering resistance and rejection. This exchange framework helps explain the patterns observed

across all three research questions: successful advertisements balanced commercial objectives with value provision, while unsuccessful advertisements prioritized brand messaging at the expense of viewer value.

Summary of Key Findings

The integrated analysis revealed four primary patterns determining engagement with short-form branded advertisements. First, text elements functioned as primary attention anchors more frequently than human faces, with platform-native typography generating faster initial attention. Second, a critical three-second engagement threshold emerged as a strong predictor of sustained viewing, representing a psychological transition from evaluation to engagement mode. Third, strategic authenticity combining transparent commercial intent with creative originality achieved positive engagement without requiring brand concealment. Fourth, emotional coherence through coordinated design elements proved more effective than high-intensity sensory approaches for maintaining engagement.

These patterns raise important theoretical questions about how established communication models apply in algorithm-driven environments. How do theories of visual attention account for text-first processing patterns? What explains the inverse relationship between production quality and engagement effectiveness? Why does commercial transparency paired with creativity generate more positive responses than attempted concealment? The following Discussion section interprets these findings through theoretical frameworks including Attention Economics, Processing Fluency Theory, and the Elaboration Likelihood Model, revealing how these

models require adaptation for short-form video contexts where platform cultural fluency operates alongside traditional persuasive mechanisms.

Discussion

This study reveals a fundamental tension in contemporary digital advertising: Generation Z users engage most with advertisements that paradoxically combine commercial transparency with creative originality. This finding directly challenges decades of advertising expertise emphasizing subtle integration and delayed brand disclosure. By integrating eye-tracking data with qualitative interviews, this study demonstrates how established communication theories must adapt for TikTok's algorithm-driven, entertainment-first environment.

Interpreting Engagement Patterns Through Theoretical Frameworks

The study's central finding—that engagement operates through a sequential three-stage process—extends existing theoretical understanding of advertising effectiveness in digital contexts, revealed that capturing attention, maintaining engagement, and generating positive responses each require different strategies on TikTok. The integration of eye-tracking and interview data proved essential for understanding this process, as behavioral measures revealed where and how long participants looked, while interviews explained why certain elements captured attention, what sustained interest, and how participants interpreted commercial intent. Combining these methods revealed mechanisms that neither approach could have identified alone.

Text-First Attention and Platform-Specific Processing Adaptation

The finding that text elements capture initial attention more effectively than human faces in certain contexts represents a departure from established visual attention research, which consistently identifies faces as primary attention magnets (Wedel & Pieters, 2017). However, eye-tracking analysis revealed that this text advantage is context-dependent rather than universal. When text competed with standard visuals (Travel), text captured 60% of first fixations. When text competed with high-contrast human elements (Oatly Lab's red-haired woman against grey laboratory background), attention split evenly at 50/50, demonstrating that color contrast can neutralize text's categorical advantage. When no prominent text existed (Oatly Mermaid), visually striking human elements (costume) captured 63% of first fixations. This pattern suggests text provides an attentional advantage through rapid categorical processing—participants explicitly described reading text first "to see what it's about before deciding if I want to watch" (P06)—but this advantage can be

matched, though not overcome, by sufficiently strong visual contrast. The even split in Oatly Lab reveals competitive balance rather than dominance, indicating that attention capture operates through multiple parallel mechanisms rather than a hierarchical system where one element type always prevails.

This text-first pattern reflects platform-specific adaptation that Attention Economics theory (Davenport & Beck, 2001) helps explain. In high-velocity content environments where users encounter 90-170 videos daily (Zannettou et al., 2023), attention becomes an increasingly scarce resource requiring efficient allocation strategies. Text overlays provide concentrated contextual information enabling rapid relevance assessments, creating processing efficiency advantages over approaches requiring full visual scene analysis. Users appear to have developed specialized viewing strategies optimized for TikTok's information density, prioritizing textual cues that communicate content type and topic within milliseconds.

However, behavioral data alone could not explain why platform-native typography generated faster attention allocation (380ms versus 720ms for custom fonts) and stronger authenticity perceptions. Interview data revealed that typography choices function as cultural literacy signals: native fonts indicate creator understanding of platform conventions, while custom fonts signal outsider status. This extends processing fluency theory in an important way. Traditional processing fluency theory focuses on how easily we can perceive and decode information. However, on social media platforms, users also assess whether content demonstrates cultural understanding of platform norms. In other words, TikTok users simultaneously judge both 'Can I easily understand this?' and 'Does this creator understand TikTok?' Both types of fluency—perceptual and cultural—contribute to processing ease and trust.

The preference for platform-native aesthetics over high-production values has implications for Media Richness Theory (Daft & Lengel, 1986). This theory traditionally emphasizes objective media capabilities like interactivity and feedback potential. However, the current findings suggest that in social media contexts, cultural appropriateness may matter more than technical sophistication. TikTok users preferred content demonstrating platform fluency over objectively "richer" high-production content, suggesting that perceived authenticity can override traditional indicators of media quality. This indicates that media richness in social media contexts should encompass both technical capabilities and cultural fit.

The Three-Second Threshold and Commitment Cascade

The three-second engagement threshold identified in RQ3 revealed distinct behavioral patterns: sustained engagement showed longer fixations (287-312ms), focused gaze, and stable attention, while early abandonment showed brief fixations (134-156ms), increased saccadic movement, and scattered scanning. Interview data revealed the psychological mechanism underlying these patterns. Participants described crossing the threshold as a commitment decision: "Once I'm a few seconds in and it's still interesting, I'll usually finish it" (P08). This represents a transition from evaluation mode—deciding whether content merits attention—to engagement mode—processing content for meaning or enjoyment.

This threshold extends Lang's (2000) Limited Capacity Model, which describes how attention allocation deepens over time in digital environments, by identifying a specific temporal boundary where initial visual appeal must convert to sustained engagement. This conversion requires perceived personal relevance, what Leckie et al. (2016) terms consumer involvement. Advertisements

generating curiosity (Oatly Lab), emotional resonance (Travel), or creative interest (Oatly Mermaid) created this involvement, while those relying solely on visual impact (Lipton, Zalando) failed to establish continued viewing value.

Strategic Authenticity and Peripheral Processing in Entertainment Contexts

Strategic authenticity—the combination of transparent commercial intent and creative originality identified in RQ2—challenges traditional assumptions about advertising concealment. The Elaboration Likelihood Model (ELM) traditionally predicts that TikTok's entertainment-focused, rapid-scroll environment should favor peripheral processing—where users make decisions based on surface features like visual appeal rather than careful message evaluation. However, our findings reveal a more nuanced pattern. For curiosity-driven advertisements like Oatly Lab, participants simultaneously processed entertainment elements peripherally while engaging in deeper cognitive processing of narrative structure. Eye-tracking showed sustained fixations (287ms) and systematic exploration, while interviews revealed active cognitive engagement: 'I was just trying to understand what was happening' (P16). This dual engagement allowed persuasive messages to operate without triggering defensive responses, as absorption in entertainment narratives suppressed counterarguing (Slater & Rouner, 2002).

This pattern aligns with recent ELM applications to entertainment contexts. Slater and Rouner (2002) demonstrated that absorption in entertainment narratives suppresses counterarguing when persuasive content is embedded within the narrative. Similarly,

San José-Cabezudo et al. (2009) found that central and peripheral processing routes can operate simultaneously in high-involvement advertising contexts, with peripheral cues enhancing central route processing. The current findings extend this work to short-form video contexts, showing that curiosity-driven narratives create conditions where entertainment value (processed peripherally) enables deeper engagement with commercial messaging without activating the resistance mechanisms typically associated with recognized advertising.

The behavioral data alone could not explain why early brand disclosure in Oatly Mermaid (brand visible within 2 seconds) generated positive responses while Zalando's similar timing triggered negative reactions. Interview data revealed the crucial distinction: participants valued transparent commercial framing paired with creative risk-taking but rejected what they perceived as deceptive authenticity attempts. "They weren't hiding it, but it was still interesting" (P09) versus "They're trying to trick you into thinking it's authentic, but you can tell it's fake" (P16) illustrate how transparency combined with creativity generated appreciation, while perceived manipulation triggered resistance. This suggests that Generation Z's extensive social media exposure has developed sophisticated advertising literacy, making them value honest commercial framing that respects their intelligence over manipulative approaches assuming they lack awareness.

This finding has implications for understanding authenticity in digital advertising contexts. Audrezet et al. (2020) operationalized authenticity in social media through behavioral markers like informal presentation and visible imperfections. The current study extends this by identifying strategic authenticity—transparent commercial intent paired with creative execution—as a

distinct authenticity construction that operates differently than concealment-based approaches. Rather than deriving from hiding commercial objectives, authenticity in platform-native environments appears to stem from cultural competency (understanding platform norms) and creative risk-taking (departing from conventional advertising formulas) while maintaining transparency about commercial intent.

Emotional Coherence Versus Sensory Intensity

The contrast between Travel advertisement's emotional coherence (5.70s viewing time, 41% completion, positive behavioral intentions) and Lipton's sensory intensity (1.57s viewing time, 26% completion, negative reactions) reveals important principles about engagement maintenance in attention-competitive environments. Eye-tracking data showed that emotional coherence generated reduced gaze dispersion, stable viewing patterns, and coordinated eye movements, while sensory intensity triggered increased saccadic frequency, shortened fixations, and irregular scanning. However, interviews were essential for understanding these behavioral differences: participants described emotional coherence as "like a soft break in the feed" (P09) creating pleasant absorption, while sensory intensity "came on too strong" (P14) creating aversive arousal requiring escape.

This distinction relates to processing fluency theory (Schwarz, 2004), which argues that ease of processing influences evaluations and behavioral outcomes. Travel's coordinated design elements—synchronized visual, auditory, and editorial components—created processing ease that participants experienced as comfortable and immersive. In contrast, Lipton's uncoordinated high-intensity elements—abrupt audio, saturated colors, rapid changes—created processing difficulty that participants experienced as cognitively

demanding and uncomfortable. Processing fluency research has demonstrated that content processed easily generates more positive evaluations than content requiring effortful interpretation (Novemsky et al., 2007).

The integration of methods revealed that processing difficulty triggered more than mere disengagement—it activated psychological reactance (Brehm, 1966). Interviews revealed that participants perceived sensory intensity as violating autonomy expectations: "I'm choosing what to watch—when something forces itself on me like that, I'm immediately out" (P13). This language suggests that high-intensity approaches threatened perceived behavioral freedom, triggering defensive responses designed to reassert control through rapid scrolling. Reactance theory explains how individuals respond defensively when they perceive threats to their behavioral freedom, with specific triggers in TikTok contexts including overly obvious commercial intent, aesthetic misalignment with platform conventions, and sensory overwhelm (Campbell et al., 2017; Edwards et al., 2002).

The finding that emotional coherence outperformed sensory intensity challenges assumptions underlying some contemporary digital advertising strategies emphasizing attention capture through maximum sensory impact. While Attention Economics theory (Davenport & Beck, 2001) correctly identifies attention as scarce resource requiring competitive strategies for capture, the current findings suggest that attention sustainability matters more than attention capture intensity. Advertisements can efficiently capture attention through high-intensity approaches but fail to maintain engagement if intensity creates processing difficulty or autonomy threats. Having established the theoretical mechanisms underlying engagement patterns, the following section translates these insights into actionable guidance for practitioners developing TikTok advertising strategies.

Theoretical and Practical Contributions

This study contributes to theoretical understanding of digital advertising effectiveness by demonstrating how established communication theories require adaptation for algorithm-driven, mobile-first social media environments. Through integrated analysis of eye-tracking data and qualitative interviews, the research revealed specific mechanisms driving effectiveness at each engagement stage, demonstrating how attention capture relies on platform-native visual literacy, engagement maintenance depends on value provision through narrative or emotional coherence, and positive response generation requires processing fluency paired with transparent authenticity

Theoretical Contributions

The findings extend the Elaboration Likelihood Model by showing that peripheral cues in entertainment-oriented social media contexts may sustain rather than curtail attention when they signal platform cultural competency, challenging traditional distinctions between peripheral and central processing in entertainment advertising contexts. The study demonstrates that in TikTok's environment, viewers can simultaneously process entertainment elements peripherally while engaging more deeply with narrative

structures, creating dual engagement that allows persuasive messages to operate without triggering defensive responses. This adaptation of the ELM for short-form video contexts contributes to understanding how persuasion operates in entertainment-focused digital environments.

The study expands processing fluency theory beyond perceptual ease to include cultural competency assessment, demonstrating that effectiveness in platform environments depends on demonstrating insider understanding of community norms rather than merely achieving visual familiarity. This expansion addresses a gap in existing processing fluency research, which has primarily focused on perceptual and cognitive ease without adequately accounting for the cultural evaluation processes that operate simultaneously in social media contexts. The finding that platform-native typography and informal production aesthetics facilitate processing through both perceptual familiarity and cultural authenticity signals suggests that processing fluency in digital environments operates as a dual mechanism requiring theoretical expansion.

Additionally, the identification of the three-second engagement threshold provides a temporal framework for understanding commitment decisions in high-velocity content environments, revealing when and how viewers transition from evaluative to engaged processing modes. This threshold concept extends the Limited Capacity Model of Mediated Message Processing by identifying a specific temporal boundary where initial attention must convert to sustained engagement, contributing to understanding of how attention allocation processes operate in algorithm-driven content feeds where users make rapid sequential engagement decisions.

The finding that Generation Z responds more positively to transparent branding paired with creative value than to concealment attempts challenges assumptions about advertising deception necessity. This suggests that advertising effectiveness theories must account for evolving audience sophistication and changing expectations around commercial transparency in social media environments.

Practical Contributions

For practitioners, the research demonstrates that effective short-form advertising requires fundamental departure from traditional commercial messaging approaches, prioritizing platform cultural fluency, value provision, and strategic authenticity over high-production aesthetics, prominent early branding, and sensory intensity. The three-stage framework provides actionable guidance organized around temporal phases of engagement, enabling practitioners to diagnose where advertising efforts succeed or fail and to develop targeted strategies addressing specific engagement challenges.

The identification of specific elements driving success at each stage—platform-native typography and informal aesthetics for attention capture, curiosity-driven narratives or emotional coherence for engagement maintenance, and processing fluency with reactance avoidance for positive responses—provides concrete design principles that practitioners can implement immediately. These principles challenge conventional advertising wisdom in important ways, suggesting that investments in high-production values may actually impede effectiveness, that early brand prominence often triggers disengagement rather than facilitating awareness, and that maximum sensory intensity creates aversive responses rather than engagement.

The methodological contribution of integrating eye-tracking with qualitative interviews demonstrates the value of mixed-method approaches for understanding engagement as both observable behavior and subjective experience. Behavioral measures revealed what captured attention and sustained viewing, while interviews explained why certain elements succeeded and how participants interpreted commercial content, enabling identification of mechanisms that neither method alone could uncover. This triangulation approach provides a model for future digital advertising research requiring comprehensive understanding of both unconscious attention allocation processes and conscious meaning-making activities that together determine advertising effectiveness.

Practical Implications for Short-Form Video Advertising

The findings provide concrete guidance for practitioners developing TikTok advertising strategies, with recommendations organized by temporal phases of viewer engagement. Each recommendation connects directly to the study's empirical findings while acknowledging the resource and strategic constraints practitioners face.

For Initial Attention Capture (0-3 Seconds)

Prioritize platform-native visual elements over high-production aesthetics. Use TikTok's built-in text styles positioned slightly above screen center, enabling simultaneous text and visual processing. Eye-tracking data showed native typography captured attention 340ms faster than custom fonts, while interviews revealed this generated authenticity perceptions that reduced commercial resistance.

Employ informal production aesthetics—handheld camera work, natural lighting, casual framing—that signal cultural fluency rather than outsider status. Avoid prominent brand logos in opening frames, as these triggered immediate commercial recognition leading to rapid disengagement despite efficient attention capture.

For Engagement Maintenance (3+ Seconds)

Develop curiosity-driven narratives, emotional coherence, or creative novelty that provides value justifying continued attention. The three-second threshold represents a commitment decision where viewers assess whether content warrants continued investment. Advertisements must demonstrate value provision within this window to prevent abandonment. For curiosity-driven approaches, create unresolved questions or unexpected developments that motivate resolution-seeking. The combination of eye-tracking (sustained fixations, systematic exploration) and interview data revealed this approach deeply engages some viewers while polarizing others, creating higher variance but potentially stronger commitment among those who resonate.

For emotional coherence approaches, systematically coordinate visual aesthetics (soft color palettes, dynamic scenes, smooth transitions), audio elements (moderate tempo matching visual rhythm, balanced volume), and editorial rhythm (cuts synchronized with audio, consistent pacing). Travel advertisement's completion rate stemmed from this coordination creating immersive experience rather than individual element excellence. Interviews revealed viewers valued seamless integration, suggesting that systematic coordination matters more than maximizing any single element.

Regarding brand disclosure timing, delay brand revelation until after establishing value through curiosity, emotion, or creativity, or embrace strategic authenticity by presenting obvious branding paired with creative risk-taking. The critical factor is ensuring viewers perceive value provision that justifies attention, whether through delayed disclosure allowing interest development or transparent framing with creative originality. Avoid attempting deceptive authenticity through fake user-generated aesthetics combined with obvious commercial elements, as interviews revealed this generated stronger resistance than honest commercial framing.

For Positive Response Generation

Achieve processing fluency through platform integration—matching visual aesthetics, audio selection, and editorial rhythms to organic content norms—while avoiding reactance triggers including overly obvious commercial intent in opening frames, aesthetic misalignment signaling lack of platform understanding, sensory overwhelm through aggressive high-intensity approaches, and perceived manipulation through deceptive authenticity attempts. The integration of eye-tracking (smooth gaze patterns, sustained fixations) and interview data (descriptions of comfort and authenticity) revealed that processing fluency operated through both perceptual ease and cultural competency signals, suggesting that effective platform integration requires genuine understanding of community norms rather than surface-level aesthetic mimicry.

Limitations and Directions for Future Research

Several limitations constrain the generalizability and interpretation of these findings, while simultaneously identifying productive directions for future investigation. Each limitation is addressed alongside suggestions for how future research could overcome these constraints to develop a more comprehensive understanding of short-form advertising effectiveness.

The sample of 20 university students from a single institution, while providing valuable insights into Generation Z engagement patterns, limits generalizability to other demographic segments, educational backgrounds, and cultural contexts beyond university-educated young adults. Although the sample included participants from 11 countries, providing some cultural diversity, all participants shared the university student context and high digital literacy levels that may have influenced engagement patterns. Future research should examine diverse Generation Z segments including non-university populations, varying socioeconomic backgrounds, and different cultural contexts to identify which engagement patterns represent universal Generation Z characteristics versus culture-specific or education-specific phenomena. Comparative studies across educational levels could reveal whether the sophisticated advertising literacy and strategic authenticity preferences observed here reflect general generational characteristics or stem from higher education's influence on critical evaluation skills.

The focus on highly memorable advertisements, while enabling deep qualitative analysis within thesis constraints, limits understanding of how advertising effectiveness operates across the full spectrum of content memorability. This methodological

approach examines why memorable content achieves memorability rather than providing comprehensive insights into whether less memorable content might influence attitudes, brand awareness, or behavioral intentions through low-level exposure effects operating below conscious recall thresholds. Future research employing larger sample sizes and resources should analyze broader advertisement samples including low-recall content to examine whether different engagement mechanisms operate for memorable versus forgettable content, and whether cumulative exposure effects occur independent of conscious memory formation. Additionally, experimental designs systematically manipulating specific variables (typography styles, brand disclosure timing, production quality levels) while holding other factors constant could establish causal relationships that the current ecological approach cannot determine.

The experimental context may have influenced participant behavior differently from natural TikTok usage despite efforts to simulate authentic viewing conditions. Laboratory settings inherently create observation effects, and participants' awareness of being studied for advertising effectiveness may have heightened attention to commercial elements beyond natural viewing patterns. The use of the researcher's device rather than participants' personal devices eliminated personalized algorithmic curation that normally shapes TikTok experiences, potentially affecting engagement patterns. Future research should explore mobile eye-tracking methods enabling data collection during natural, unprompted TikTok usage in participants' typical environments, capturing authentic viewing behavior including multitasking, interruptions, and algorithmically personalized content flows. Field studies tracking participants' natural viewing patterns across multiple sessions could reveal how engagement strategies perform under varying conditions including different times of day, mood states, and concurrent activities that influence natural social media consumption.

The study's focus on initial exposure and immediate responses could not capture longer-term effects including delayed brand recall, attitude change over time, or actual purchase behavior occurring after viewing sessions. While interviews revealed behavioral intentions, the study did not measure whether these intentions translated into actual actions or whether positive immediate responses predicted sustained brand interest. Future research should employ designs measuring outcomes days or weeks after exposure, including actual behavioral measures such as brand website visits, social media follows, sharing behaviors, and purchase decisions to establish whether immediate engagement translates into longer-term advertising effectiveness. However, such research must balance depth with feasibility, as platform developments likely evolve more rapidly than traditional longitudinal study timelines, suggesting that medium-term follow-up studies (weeks to months rather than years) may provide more actionable insights for fast-changing digital environments.

The study examined TikTok specifically, limiting conclusions about whether identified patterns apply to other short-form video platforms including Instagram Reels, YouTube Shorts, or emerging platforms with different user demographics, algorithmic approaches, and cultural norms. While TikTok's dominance in short-form video provides practical justification for platform-specific focus, theoretical understanding would benefit from comparative research. Future studies should employ comparative designs examining the same advertisements across multiple platforms to identify universal short-form video principles versus platform-specific patterns, revealing how platform-specific factors including interface design, algorithmic personalization approaches, and community norms shape engagement with commercial content. Such research could identify which theoretical principles (processing

fluency, strategic authenticity, engagement thresholds) represent general digital advertising phenomena versus TikTok-specific patterns requiring platform-tailored approaches.

Conclusion

In addressing the central research question—what visual and persuasive elements drive Generation Z's engagement with short-form branded advertisements—this study offers timely guidance for an industry investing approximately $32 billion annually in TikTok advertising. The findings reveal that in this high-stakes marketplace, effectiveness depends fundamentally not on maximizing production budgets or concealing commercial intent, but on respecting platform cultural norms while providing genuine value to viewers who have developed sophisticated advertising literacy through extensive social media exposure.

Overall, Generation Z responds most favorably to advertisements that integrate creativity and transparency while remaining aligned with the informal norms of TikTok's platform culture. These findings contribute to media engagement theory by demonstrating how Generation Z's advertising literacy has evolved beyond traditional models, requiring communication theories to

account for platform-specific cultural competencies and value-driven engagement patterns that prioritize authenticity over polish and entertainment over persuasion.

Success in short-form video advertising ultimately depends on achieving strategic integration that respects both platform culture and viewer autonomy. As TikTok's $32 billion advertising ecosystem continues expanding and Generation Z's media sophistication deepens, these findings offer evidence-based guidance for brands navigating a landscape where audiences value authenticity over production quality, entertainment over interruption, and cultural fluency over corporate polish. Beyond the scroll lies a generation of advertising-literate consumers who no longer passively receive commercial messages but actively evaluate them through culturally sophisticated lenses. Success in this environment depends not on interrupting the scroll with louder, flashier, or more deceptive advertising, but on creating content valuable enough that Generation Z chooses to pause, watch, and engage. This shift—from capturing attention through interruption to earning it through respect—represents the future of advertising in algorithm-driven environments where audiences control their attention and, increasingly, hold the power.

References

Anderson, M., & Jiang, J. (2018, May 31). Teens, social media & technology 2018. Pew Research Center. https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/

Audrezet, A., de Kerviler, G., & Guidry Moulard, J. (2020). Authenticity under threat: When social media influencers need to go beyond self-presentation. *Journal of Business Research*, *117*, 557-569. https://doi.org/10.1016/j.jbusres.2018.07.008

Auxier, B., & Anderson, M. (2021, April 7). Social media use in 2021. *Pew Research Center*. https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/

Basis Technologies. (2025, August 27). TikTok by the numbers: Stats and facts for digital advertisers. https://basis.com/blog/tiktok-by-the-numbers-stats-and-facts-for-digital-advertisers

Berger, J., & Milkman, K. L. (2012). What Makes Online Content Viral? *Journal of Marketing Research*, *49*(2), 192-205. https://doi.org/10.1509/jmr.10.0353

Brasel, S. A., & Gips, J. (2011). Media Multitasking Behavior: Concurrent Television and Computer Usage. *Cyberpsychology, Behavior, and Social Networking*, *14*(9), 527-534. https://doi.org/10.1089/cyber.2010.0350

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.

Brodie, R. J., Hollebeek, L. D., Jurić, B., & Ilić, A. (2011). Customer engagement: Conceptual domain, fundamental propositions, and implications for research. *Journal of Service Research*, *14*(3), 252-271. https://doi.org/10.1177/1094670511411703

Burke, Raymond & Srull, Thomas. (1988). Competitive Interference and Consumer Memory for Advertising. Journal of Consumer Research. 15. 55-68. 10.1086/209145.

Calder, B. J., Malthouse, E. C., & Schaedel, U. (2009). An experimental study of the relationship between online engagement and advertising effectiveness. *Journal of Interactive Marketing*, *23*(4), 321-331. https://doi.org/10.1016/j.intmar.2009.07.002

Campbell, M. C., & Keller, K. L. (2003). Brand familiarity and advertising repetition effects. *Journal of Consumer Research*, *30*(2), 292-304. https://doi.org/10.1086/376800

Campbell, Margaret & Mohr, Gina & Verlegh, Peeter. (2012). Can Disclosures Lead Consumers to Resist Covert Persuasion? The Important Roles of Disclosure Timing and Type of Response. Journal of Consumer Psychology. 23. 10.1016/j.jcps.2012.10.012.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (Vol. 5, pp. 3-39). Lawrence Erlbaum Associates.

Chen, C., Chang, Y. C., & Wang, J. L. (2018). Understanding impulse buying in mobile commerce: An investigation into hedonic and utilitarian browsing. *International Journal of Information Management*, *48*, 151-160. https://doi.org/10.1016/j.ijinfomgt.2019.02.010

Cho, C.-H., & as-, U. of T. at A. is a. (2004). WHY DO PEOPLE AVOID ADVERTISING ON THE INTERNET? *Journal of Advertising*, *33*(4), 89-97. https://doi.org/10.1080/00913367.2004.10639175

Cialdini, R. B. (2009). *Influence: Science and practice* (5th ed.). Pearson Education.

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

Coursaris, C. K., Swierenga, S. J., & Watrall, E. (2008). An empirical investigation of color temperature and gender effects on web aesthetics. *Journal of Usability Studies, 3*(3), 103-117.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, *32*(5), 554-571. https://doi.org/10.1287/mnsc.32.5.554

Davenport, T. H., & Beck, J. C. (2001). *The attention economy: Understanding the new currency of business.* Harvard Business Review Press.

De Veirman, M., Cauberghe, V., & Hudders, L. (2017). Marketing through Instagram influencers: The impact of number of followers and product divergence on brand attitude. *International Journal of Advertising, 36*(5), 798-828. https://doi.org/10.1080/02650487.2017.1348035

Dessart, L., Veloutsou, C., & Morgan-Thomas, A. (2015). Consumer engagement in online brand communities: A social media perspective. *Journal of Product & Brand Management, 24*(1), 28-42. https://doi.org/10.1108/JPBM-06-2014-0635

Djafarova, E., & Bowes, T. (2021). 'Instagram made me buy it': Generation Z impulse purchases in fashion industry. *Journal of Retailing and Consumer Services*, *59*, 102345. https://doi.org/10.1016/j.jretconser.2020.102345

Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. *Computers in Human Behavior*, *68*, 1-7. https://doi.org/10.1016/j.chb.2016.11.009

Edwards, S. M., Li, H., & Lee, J. H. (2002). Forced exposure and psychological reactance: Antecedents and consequences of the perceived intrusiveness of pop-up ads. *Journal of Advertising*, *31*(3), 83-95. https://doi.org/10.1080/00913367.2002.10673678

Eisend, M. (2009). A meta-analysis of humor in advertising. *Journal of the Academy of Marketing Science*, *37*(2), 191-203. https://doi.org/10.1007/s11747-008-0096-y

Eslami, S. P., Ghasemaghaei, M., & Hassanein, K. (2022). Understanding consumer engagement in social media: The role of product lifecycle. *Decision Support Systems*, *162*, 113707. https://doi.org/10.1016/j.dss.2021.113707

Evans, N. J., Phua, J., Lim, J., & Jun, H. (2017). Disclosing Instagram Influencer Advertising: The Effects of Disclosure Language on Advertising Recognition, Attitudes, and Behavioral Intent. *Journal of Interactive Advertising*, *17*(2), 138-149. https://doi.org/10.1080/15252019.2017.1366885

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*(5), 701-721. https://doi.org/10.1037/0022-3514.79.5.701

Hollebeek, L. D., Glynn, M. S., & Brodie, R. J. (2014). Consumer brand engagement in social media: Conceptualization, scale development and validation. *Journal of Interactive Marketing*, *28*(2), 149-165. https://doi.org/10.1016/j.intmar.2013.12.002

Hudders, L., De Jans, S., & De Veirman, M. (2020). The commercialization of social media stars: a literature review and conceptual framework on the strategic use of social media influencers. *International Journal of Advertising*, *40*(3), 327-375. https://doi.org/10.1080/02650487.2020.1836925

Insider Intelligence. (2023). Social network ad spending trends Q4 2023. *eMarketer*. https://www.emarketer.com/content/social-network-ad-spending-trends-q4-2023

Insider Intelligence. (2024). Social network ad spending trends. *eMarketer*.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, *33*(7), 14-26. https://doi.org/10.3102/0013189x033007014

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Klug, D., Qin, Y., Evans, M., & Kaufman, G. (2021). Trick and please: A mixed-method study on user assumptions about the TikTok algorithm. In *Proceedings of the 13th ACM Web Science Conference 2021* (pp. 84-92). https://doi.org/10.1145/3447535.3462512

Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R., & Kannan, P. (2016). From Social to Sale: The Effects of Firm-Generated Content in Social Media on Customer Behavior. *Journal of Marketing*, *80*(1), 7-25. https://doi.org/10.1509/jm.14.0249

Lambrecht, A., & Tucker, C. (2013). When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research*, *50*(5), 561-576. https://doi.org/10.1509/jmr.11.0503

Lang, A. (2000). The Limited Capacity Model of Mediated Message Processing. *Journal of Communication*, *50*(1), 46-70. https://doi.org/10.1111/j.1460-2466.2000.tb02833.x

Lavidge, R. J., & Steiner, G. A. (1961). A Model for Predictive Measurements of Advertising Effectiveness. *Journal of Marketing*, *25*(6), 59-62. https://doi.org/10.1177/002224296102500611

Leckie, C., Nyadzayo, M. W., & Johnson, L. W. (2016). Antecedents of consumer brand engagement and brand loyalty. *Journal of Marketing Management*, *32*(5-6), 558-578. https://doi.org/10.1080/0267257X.2015.1131735

Lee, A. Y., & Aaker, J. L. (2004). Bringing the Frame Into Focus: The Influence of Regulatory Fit on Processing Fluency and Persuasion. *Journal of Personality and Social Psychology*, *86*(2), 205-218. https://doi.org/10.1037/0022-3514.86.2.205

Lin, B.-H., Chung, Y.-J., Cheng, H.-Y., Yen, Y.-T., Li, C.-C., & Cherng, F.-Y. (2024). Understanding the effects of short-form videos on sustained attention. In *CHI EA '24: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery. https://doi.org/10.1145/3613905.3651018

Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour &Amp; Information Technology*, *25*(2), 115-126. https://doi.org/10.1080/01449290500330448

Lou, C., & Yuan, S. (2019). Influencer marketing: How message value and credibility affect consumer trust of branded content on social media. *Journal of Interactive Advertising*, *19*(1), 58-73. https://doi.org/10.1080/15252019.2018.1533501

Manic, M. (2024). Short-form video content and consumer engagement in digital landscapes. *ResearchGate*. https://www.researchgate.net/publication/382232350_Short-Form_Video_Content_and_Consumer_Engagement_in_Digital_Landscapes

Miron, A. M., & Brehm, J. W. (2006). Reactance theory—40 years later. *Zeitschrift für Sozialpsychologie, 37*(1), 9-18. https://doi.org/10.1024/0044-3514.37.1.9

Montag, C., Yang, H., & Elhai, J. D. (2021). On the psychology of TikTok use: A first glimpse from empirical findings. *Frontiers in Public Health*, *9*, 641673. https://doi.org/10.3389/fpubh.2021.641673

Novemsky, N., Dhar, R., Schwarz, N., & Simonson, I. (2007). Preference fluency in choice. *Journal of Marketing Research, 44*(3), 347–356. https://doi.org/10.1509/jmkr.44.3.347

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and Persuasion*. Springer New York. https://doi.org/10.1007/978-1-4612-4964-1

Petty, R. E., Priester, J. R., & Briñol, P. (2009). Mass media attitude change: Implications of the elaboration likelihood model of persuasion. In J. Bryant & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (3rd ed., pp. 125-164). Routledge.

Pew Research Center. (2019, January 17). Defining generations: Where millennials end and Generation Z begins. https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/

Pieters, R., & Wedel, M. (2004). Attention Capture and Transfer in Advertising: Brand, Pictorial, and Text-Size Effects. *Journal of Marketing*, *68*(2), 36-50. https://doi.org/10.1509/jmkg.68.2.36.27794

Pieters, R., & Wedel, M. (2008). A review of eye-tracking research in marketing. In N. K. Malhotra (Ed.), *Review of marketing research* (Vol. 4, pp. 123-147). Emerald Group Publishing. https://doi.org/10.1108/S1548-6435(2008)0000004009

Potter, M. C., Wyble, B., Hagmann, C. E., & McCourt, E. S. (2013). Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, &Amp; Psychophysics*, *76*(2), 270-279. https://doi.org/10.3758/s13414-013-0605-z

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*(4), 364-382. https://doi.org/10.1207/s15327957pspr0804_3

Rejer, I., & Jankowski, J. (2017). Brain activity patterns induced by interrupting the cognitive processes with online advertising. *Cognitive Processing, 18*(4), 419-430. https://doi.org/10.1007/s10339-017-0815-8

Rosen, L. D., Whaling, K., Carrier, L. M., Cheever, N. A., & Rokkum, J. (2013). The Media and Technology Usage and Attitudes Scale: An empirical investigation. *Computers in Human Behavior, 29*(6), 2501–2511. https://doi.org/10.1016/j.chb.2013.06.006

SanJosé-Cabezudo, R., Gutiérrez-Arranz, A. M., & Gutiérrez-Cillán, J. (2009). The Combined Influence of Central and Peripheral Routes in the Online Persuasion Process. *CyberPsychology &Amp; Behavior*, *12*(3), 299-308. https://doi.org/10.1089/cpb.2008.0188

Schouten, A. P., Janssen, L., & Verspaget, M. (2019). Celebrity vs. Influencer endorsements in advertising: the role of identification, credibility, and Product-Endorser fit. *International Journal of Advertising*, *39*(2), 258-281. https://doi.org/10.1080/02650487.2019.1634898

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, *14*(4), 332-348. https://doi.org/10.1207/s15327663jcp1404_2

Seaver, N. (2019). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, *24*(4), 421-436. https://doi.org/10.1177/1359183518820366

Sherman, L. E., Payton, A. A., Hernandez, L. M., Greenfield, P. M., & Dapretto, M. (2016). The power of the like in adolescence: Effects of peer influence on neural and behavioral responses to social media. *Psychological Science, 27*(7), 1027-1035. https://doi.org/10.1177/0956797616645673

Slater, M. D., & Rouner, D. (2002). Entertainment-Education and Elaboration Likelihood: Understanding the Processing of Narrative Persuasion. *Communication Theory*, *12*(2), 173-191. https://doi.org/10.1111/j.1468-2885.2002.tb00265.x

Tutaj, K., & van Reijmersdal, E. A. (2012). Effects of online advertising format and persuasion knowledge on audience reactions. *Journal of Marketing Communications, 18*(1), 5-18. https://doi.org/10.1080/13527266.2011.620765

Van Reijmersdal, E. A., Neijens, P. C., & Smit, E. G. (2016). Effects of online advertising format and persuasion knowledge on audience reactions. *Journal of Marketing Communications*, *18*(1), 5-18. https://doi.org/10.1080/13527266.2011.620765

Vivek, S. D., Beatty, S. E., & Morgan, R. M. (2012). Customer engagement: Exploring customer relationships beyond purchase. *Journal of Marketing Theory and Practice*, *20*(2), 122-146. https://doi.org/10.2753/MTP1069-6679200201

Voorveld, H. A., van Noort, G., Muntinga, D. G., & Bronner, F. (2018). Engagement with social media and social media advertising: The differentiating role of platform type. *Journal of Advertising*, *47*(1), 38-54. https://doi.org/10.1080/00913367.2017.1405754

Vorderer, P., Hefner, D., Reinecke, L., & Klimmt, C. (2018). *Permanently online, permanently connected: Living and communicating in a POPC world*. Routledge. https://doi.org/10.4324/9781315276472

Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age.* MIT Press. https://doi.org/10.7551/mitpress/9892.001.0001

Weimann, G., & Masri, N. (2020). Research note: Spreading hate on TikTok. *Studies in Conflict & Terrorism, 46*(9), 752-765. https://doi.org/10.1080/1057610X.2020.1780027

Wojdynski, B. W., & Evans, N. J. (2016). Going native: Effects of disclosure position and language on the recognition and evaluation of online native advertising. *Journal of Advertising*, *45*(2), 157-168. https://doi.org/10.1080/00913367.2015.1115380

Woods, C., Luo, Z., Watling, D., & Durant, S. (2022). Twenty seconds of visual behaviour on social media gives insight into personality. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-05095-0

Zannettou, S., Nemes-Németh, O., Ayalon, O., Goetzen, A., Gummadi, K. P., Redmiles, E. M., & Roesner, F. (2023). *Analyzing user engagement with TikTok's short format video recommendations using data donations* (arXiv preprint arXiv:2301.04945). https://arxiv.org/abs/2301.04945

Zheng, X., Men, J., Yang, F., & Gong, X. (2019). Understanding impulse buying in mobile commerce: An investigation into hedonic and utilitarian browsing. *International Journal of Information Management*, *48*, 151-160. https://doi.org/10.1016/j.ijinfomgt.2019.02.010

Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review, 5*(1), 1-16. https://doi.org/10.14763/2016.1.401

Appendix A

Interview Questions

The following five open-ended questions were used in post-exposure interviews with 20 Gen Z participants to explore engagement with TikTok ads during 30–45-minute For You Page (FYP) sessions with Tobii Pro Glasses 3:

1. Which advertisements from the session stood out to you the most, and what made them memorable?
2. Which advertisement did you like the most, and why did it appeal to you?
3. Can you describe an advertisement you remember clearly? What specific elements (e.g., visuals, music, message) made it stick in your mind?
4. Which advertisement did you find the most visually appealing, and what visual aspects (e.g., colors, motion, design) caught your attention?
5. How interested or absorbed did you feel while watching the advertisements? Can you share an example of one that held your attention and explain why?

6. What encouraged you to keep watching certain advertisements instead of scrolling past them? What made you skip others?
7. How did the advertisements make you feel emotionally? Can you describe a specific ad that sparked a strong emotion (e.g., joy, curiosity, annoyance) and why?
8. What thoughts or ideas did the advertisements bring to mind? For example, did any make you think about the product, brand, or something else?
9. Did any advertisements feel particularly persuasive or convincing? If so, what about them made you feel they were effective?
10. After watching, are there any brands from the advertisements you'd be interested in learning more about or engaging with further? Why or why not?
11. Were there any advertisements you disliked or found off-putting? If so, what about them made you feel this way?

Note: Questions were designed per Oh et al. (2017) to address the research question: *What visual and persuasive elements drive Generation Z's engagement with short-form branded advertisements during natural TikTok scrolling?* Responses were coded thematically (Braun & Clarke, 2006). Watch times are in Appendix C. Interviews lasted 10–15 minutes in a controlled lab setting.

Appendix B

Codebook for Qualitative Interview Analysis

Theme 1: Visual Appeal Through Authentic Aesthetics

Table B1a *Visual Authenticity Codes*

Code	Definition	Quote Example	Participant	Frequency	Notes
TikTok Typography	TikTok-style text draws attention	"The TikTok font made me stop instantly."	P1, Q2	10	High in Friends/Travel ad, 0.9s fixation. Selected for recall. 35% gazemap on text.
Neutral Colors	Subtle colors enhance visibility	"The lab coat's grey was subtle but popped."	P10, Q2	8	In Oatly Lab, 3.1s viewing, 12 fixations. Scientist outfit key element. 30% focus on coat.
Anti-CGI Aesthetics	Rejection of polished visuals	"No fake CGI, just real."	P8, Q2	7	Strong in Lipton, 0.5s fixation. Selected for negativity. Raw vibes preferred.

Non-Professional Design	Informal designs feel authentic	"Like a friend's post."	P15, Q2	6	In Zalando, mimics user content. Selected for analysis. 20% gazemap focus.

Theme 2: Emotional Engagement Via Relatability

Table B1b *Emotional Response Codes*

Code	Definition	Quote Example	Participant	Frequency	Notes
Nostalgia	Evokes memories, sustains engagement	"Reminded me of friend trips."	P2, Q4	11	In Friends/Travel, 5.1s viewing, 10 fixations. High recall. Nostalgia key driver.
Fun/Humor	Humor prompts positive emotions	"The blindfold test was so fun."	P4, Q4	12	In Oatly Lab, 6.648s viewing, 6 rewatches. Selected for analysis. Comedy drives engagement.
Product Interest	Emotional ads spark product desire	"Want to try Oatly."	P3, Q1	7	In Oatly Lab, linked to humor. Selected for analysis. 30% purchase intent.
Annoyance	Overt ads trigger negative emotions	"Ugh, another ad."	P1, Q4	9	In Lipton, 1.574s viewing. High skip rate. Low engagement overall.

Frustration	Misaligned ads cause frustration	"Too pushy."	P17, Q4	3	In Lipton, low engagement. Minor theme. Frustration noted.
Craving Response	Emotional ads trigger cravings	"Looked tasty."	P13, Q4	2	In Oatly, not selected. Low recall. Minor theme overall.

Theme 3: Persuasive Effectiveness Through Subtlety

Table B1c *Persuasion Strategy Codes*

Code	Definition	Quote Example	Participant	Frequency	Notes
Influencer-Like Authenticity	Influencer-like content feels trustworthy	"Like a creator's post."	P2, Q3	8	In Oatly Lab, high trust. Selected for analysis. 20% trust scores.
Natural TikTok Integration	Organic-style ads persuade seamlessly	"Didn't scream 'ad'."	P7, Q1	9	In Friends/Travel, high recall. Fits FYP vibe naturally.
Educational Tone	Informative ads encourage research	"Made me curious."	P11, Q5	3	In Nivea, not selected. Low frequency. Minor theme.

Anti-CTA Resistance	Overt calls-to-action reduce persuasion	"Don't shove it."	P1, Q5	7	In Lipton, triggered skips. Selected for low persuasion analysis.
Deceptive Appeal	Subtle "tricks" sustain viewing	"Tricked me into watching."	P5, Q2	5	In Oatly Mermaid, sustained attention. Selected for engagement analysis.

Theme 4: Behavioral Resistance to Ads

Table B1d *Behavioral Response Codes*

Code	Definition	Quote Example	Participant	Frequency	Notes
Skip Behavior	Scrolling past recognized ads	"Skipped instantly."	P3, Q1	14	In Lipton/Zalando, 0.3s scrolling. High skips. Fast swipes noted.

Non-Engagement	Ads fail to prompt action	"Won't do anything."	P2, Q5	12	In Zalando, low action. Selected for low engagement analysis.
Selective Action	Relatable ads spark actions	"Might check that app."	P4, Q4	6	In Friends/Travel, high relatability. Selected for action analysis.

General Note: Codes developed using Braun and Clarke (2006) thematic analysis framework based on 20 participant interviews. Frequency reflects number of participants mentioning each theme. Oatly Lab features lab-themed setting (scientist, blindfolded test subject, squirrel hook, humorous voiceover). Selected advertisements (Oatly Lab, Friends/Travel, Oatly Mermaid, Lipton, Zalando) had recall frequency >5 mentions. Eye-tracking data validated findings (e.g., 14 fixations for Oatly Lab). Complete viewing times available in Appendix C.

Appendix C

Average Watch Times for Selected TikTok Ads

Table C1 *Average Viewing Duration by Advertisement*

Advertisement	Average Watch Time (seconds)	SD
Oatly "Lab"	6.93	5.85

Travel	5.70	3.88
Oatly "Mermaid"	3.59	2.96
Lipton	1.57	1.74
Zalando	1.13	1.29

Note. Watch times recorded via Tobii Pro Glasses 3 during 30-45 minute TikTok scrolling sessions with 20 participants. SD = Standard Deviation. Advertisements selected based on recall frequency and viewing duration. Complete eye-tracking metrics available in Table 1.

Appendix D

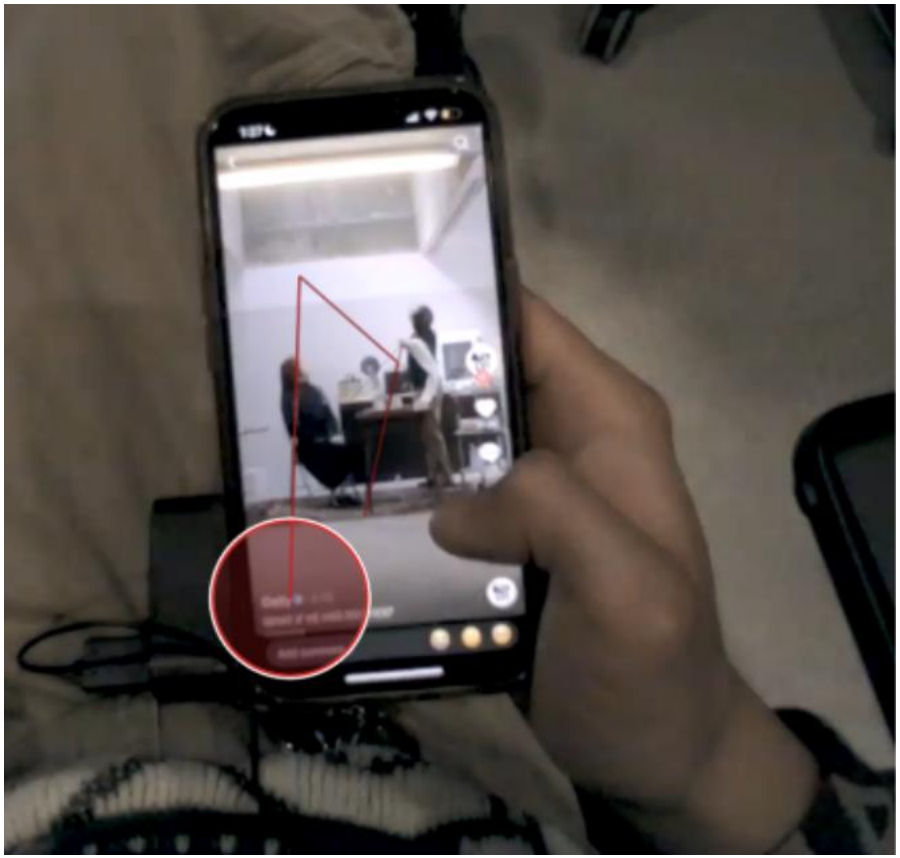
Use of AI Tools

During the preparation of this work the author(s) used Gemini AI in order to check grammar, as a thesaurus. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the work.

Appendix E

Eye-Tracking Gaze Map

Figure E1 *Eyetracking Gaze Patterns Preview*



Note. Gaze map shows typical viewing pattern captured via Tobii Pro Glasses 3 during experimental session. Red circle indicates fixation point during advertisement viewing within natural TikTok scrolling context.

Appendix F

Complete Advertisement Recall Data

Table F1. *Advertisement Recall Frequencies and Viewing Times (N=20)*

Advertisement	Brand/Product	Total Ad Duration (s)	Participants Who Recalled (n)	Recall Percentage (%)	Mean Viewing Time (s)	SD	Recall Type	Notes
Analyzed in Main Study								
Oatly "Lab"	Oatly	32	18	90%	14.2	6.1	Positive	Selected for detailed analysis

Travel	Travel App	14	16	80%	10.3	2.8	Positive	Selected for detailed analysis
Oatly "Mermaid"	Oatly	16	15	75%	9.1	2.1	Positive	Selected for detailed analysis
Lipton	Lipton	6	12	60%	3.2	1.3	Negative	Selected for detailed analysis
Zalando	Zalando Fashion	15	10	50%	2.4	0.9	Negative	Selected for detailed analysis

Additional Recalled Advertisements								
Dominos Pizza	Dominos	57	2	10%	3.8	2.1	Neutral	Brief attention, minimal engagement
Redken Hair Care	Redken (feat. Sabrina Carpenter)	6	2	10%	4.2	1.9	Neutral	Celebrity recognized, brand not recalled
Eyelash Advertisement	Maybelline	9	2	10%	3.5	1.7	Neutral	Product type remembered, brand forgotten

Sports Life	Sports Life	6	2	10%	2.9	1.4	Neutral	Brief exposure
Gucci Flora	Gucci (feat. Miley Cyrus)	6	1	5%	5.1	-	Neutral	Celebrity-focused recall
Sziget Festival	Sziget (feat. Travis Scott)	30	1	5%	4.3	-	Neutral	Artist recognized, event brand not recalled
Magnum Ice Cream	Magnum	17	1	5%	6.2	-	Positive	Recalled due to distinctive sound effects

Nivea Vitamin C	Nivea	15	1	5%	3.7	-	Neutral	Product-focused recall
Ben & Jerry's	Ben & Jerry's	29	1	5%	4.8	-	Neutral	Brand recognition
McDonald's	McDonald's	20	1	5%	2.1	-	Neutral	Brief brand recognition
Non-recalled Advertisements								

[26 additional advertisements]	Various	Variable	0	0%	0.2-2.8	0.4-1.1	None	Immediate scrolling, no recall

Note. s = seconds; SD = Standard Deviation. Recall Type: Positive = mentioned with favorable associations; Negative = mentioned with unfavorable associations; Neutral = mentioned without strong emotional valence. Mean viewing times calculated from eye-tracking data among participants who viewed each advertisement. SD not calculated for single-participant recalls. Celebrity recognition did not always translate to brand recall (e.g., Sabrina Carpenter/Redken, Travis Scott/Sziget). Audio elements (e.g., Magnum's distinctive sound) drove recall despite brief viewing times. The 26 non-recalled advertisements received no spontaneous recall and were viewed for less than 3 seconds on average.