How Should Content Creators Narrate Their Content? The Impact of Emotionality Flips on Audience Engagement

Forthcoming in the *Journal of the Academy of Marketing Science*
January 26, 2026

Hesam Ghasemi
PhD Candidate in Marketing at ESSEC Business School
mohammadhesam.hajighasemi@essec.edu
+33 6 67 61 83 49
3 Av. Bernard Hirsch, 95000 Cergy, France

Vivek Astvansh (Corresponding author)
Associate Professor of Quantitative Marketing and Analytics at McGill University
Academic Director, Bensadoun School of Retail Management,
Adjunct Associate Professor, Department of Informatics, Luddy School of Informatics, Science, and Engineering, Indiana University Bloomington
Affiliate, Environmental Resilience Institute, Indiana University
vivek.astvansh@mcgill.ca
+1 (514) 222-4037
1001 Rue Sherbrooke O, Montreal, QC H3A 1G5, Canada

Amir Sepehri
Assistant Professor of Marketing at ESSEC Business School
sepehri@essec.edu
+1 (519) 697-8440
3 Av. Bernard Hirsch, 95000 Cergy, France

Declarations

- Ethical Approval: Not applicable
- Availability of supporting data: Available from the first author
- Competing interest: The authors declare that they have no competing interests.
- Funding: None
- Acknowledgments: Not applicable.

How Should Content Creators Narrate Their Content? The Impact of Emotionality Flips on Audience Engagement

ABSTRACT

The creator economy marks a pivotal shift in influencer marketing. Content creators create *long* content to engage their audience on social media platforms and generate revenue through engagement. We hypothesize that audience engagement with content is a function of *emotionality flips*: the number of times emotionality flips direction (from a positive to a negative slope, and vice versa) across one content unit (e.g., a sentence). Five studies—including 33,598 podcast episodes, 3,381 TED Talks, two lab experiments, and one synthetic validation—support this hypothesis. Furthermore, arousal is the underlying mechanism; that is, emotionality flips raise the audience's arousal, which in turn boosts audience engagement. Lastly, emotionality flips' positive effect is stronger for (1) content that is easier (vs. harder) to process, and (2) audiences with high (vs. low) need-for-affect. This research advances understanding of emotionality dynamics in the creator economy and offers insights for creators, platforms, and consumers.

Keywords: influencers, natural language processing, automated text analysis, narrative transportation, digital opinion leaders

INTRODUCTION

The *creator*[1] *economy* refers to "a space where individuals, known as creators, generate income through producing and sharing original content, products, or services on digital platforms" (Peres et al., 2024, p.1). It includes influencers, podcasters, video creators, bloggers, and live streamers. The creator economy is valued at $192 billion in 2025 (Kumar 2025) and projected to reach nearly half a trillion U.S. dollars by 2027 (Goldman Sachs 2023), with a cumulative average growth rate of 23% (Kumar 2025). The economy boasts 207 million creators globally (Kumar 2025), resulting in 1.5 million digital creator jobs in the United States alone in 2024 (Fischer 2025). These statistics highlight the need for research on the creator economy.

On the surface, the *creator economy* term broadens *influencer marketing*[2] in two ways. First, the prefix "creator" replaces the word "influencer" and thus overcomes the negative connotation of influencing/persuading others (Campbell and Kirmani 2000). Second, the suffix "economy" reminds stakeholders of the economic significance of the created content rather than limiting their view of it as just another marketing activity. Interestingly, this broadening of the term is also a marketing exercise, and influencer marketing has become a key catalyst in growing the creator economy (Kozinets et al. 2023).

On a deeper level, the creator economy characterizes a pivotal shift in influencer marketing. It provides anyone with internet access an unprecedented opportunity to turn their creativity into an entrepreneurial career. This growth reflects a significant shift in content production, distribution, and monetization, moving away from traditional media platforms

[1] A creator is an individual or an organization who produces content, often with the aim of generating revenue through audience engagement with the content (Kozinets et al. 2023).

[2] *Influencers* are individuals who turn their online followers into a target audience for marketing messages in exchange for compensation from firms (Mardon et al. 2023). *Influencer marketing* involves a firm selecting and incentivizing influencers to engage their followers on social media, leveraging the unique resources these influencers provide to promote the firm's offerings, with the ultimate goal of enhancing the firm's performance (Leung et al., 2022b).

such as television, where media organizations mainly generate content. By the end of 2023, there were over 1 million content creators on TikTok (Duarte 2023a), more than 23 million on Instagram (Howarth 2024), and over 60 million on YouTube (Duarte 2023b). Unlike their traditional media counterparts, social media platforms enable creators to bypass gatekeepers, such as publishers and broadcasters, engage directly with their audience, and assume greater control over their content (Gardner and Lehnert 2016). The expanding market for the creator economy has attracted more content creators, prompting them to enhance content quality and narrativity to boost audience engagement (Jiang et al. 2024). What type of content should creators produce to receive high audience engagement? The success of the creator economy may hinge on the answer to this question.

Prior research has shown that content emotionality is a central determinant of audience engagement with traditional media (Tellis et al. 2019) and social media (Berger 2014). Therefore, we use emotionality as our starting point. Next, although previous research has shown that emotionality drives engagement (e.g., Berger et al. 2023; Berger and Milkman 2012; Lee 2021; Milkman and Berger 2014), academics and practitioners have paid less attention to how and why audience engagement is a function of *emotionality flips*,[3] the number of times emotionality flips direction (from a positive slope to a negative slope or vice versa) from one unit of content (e.g., a sentence) to another (see Berger et al., 2021 for how sentiment volatility can lead to virality). This lack of attention motivates our research and is our point of departure. We extend prior research by hypothesizing that emotionality flips increase audience engagement with narrative content. We reason that longer content is more likely to regress to the mean in emotionality, as more emotional sentences will be canceled out by less emotional ones, rendering the "average" emotionality less insightful.

[3] Emotionality pivots and emotionality reversals are practical alternative terms to describe similar directional changes in the emotionality of a content.

Regression analyses of 33,598 YouTube podcast episodes published between 2019 and 2024 (Study 1) and 3,381 TED Talks published between 2006 and 2020 (Study 2) support the hypothesized association. Two lab experiments (Studies 3 and 4) provide causal evidence, and one synthetic study (Study 5) validates the robustness of the findings. The key insight is that what drives audience engagement with long-form narratives is not simply the emotional valence or intensity, but rather the number of shifts—or *flips*—in emotionality valence throughout the content. Moreover, Studies 1 and 2 demonstrate a moderating effect of narrative fluency, Study 3 clarifies that arousal is an underlying psychological mechanism, and Study 4 establishes the moderating role of consumers' Need for Affect (NFA).

Our research makes three contributions. First, it advances work on audience engagement in the creator economy (Rizzo, et al., 2023a; Rizzo, et al., 2023b; Goldenberg et al., 2023; Gu et al., 2024; Lanz et al., 2024; Leung et al., 2022; Mosley et al., 2024; see Web Appendix A, Table A1 for a summary of representative research) by showing that beyond emotional valence and intensity, changes in emotionality over the course of a narrative meaningfully shape engagement in long-form content. Second, it contributes to narrativity research (Hamby and Escalas 2024; Herzenstein et al. 2011; van Laer et al. 2014, 2019; see Web Appendix A, Table A2 for a summary of representative research) by positioning emotionality flips as a discourse-level feature (i.e., how the content is expressed; Chatman 1978; Culler 1981) that introduces emotional dynamics into messages and facilitates narrative persuasion (Argo et al., 2008; Escalas, 2007). Third, it offers practical implications: creators can use emotionality flips to enhance engagement, platforms may integrate them into recommendation systems, and consumers can leverage emotional variability to enrich affective experiences.

NARRATIVITY IN CONTENT

Content creators can persuade the audience through *narrative transportation*—"the degree to which (1) a consumer empathizes with the story's characters and (2) the plot stimulates consumers' imagination, causing them to temporarily experience a suspension of reality during the story" (van Laer et al., 2014, p. 799-800; see also Escalas 2007; Green and Brock 2000; Hamby and Brinberg 2016). By immersing themselves in a story world (Green and Brock 2000), audiences tend to process the content less critically (Adaval and Wyer Jr. 1998; Escalas 2007) and sense a real-life experience (Green and Brock 2000).

The level of narrativity influences the degree to which audiences experience narrative transportation and become immersed in the story world (Nell, 1988). Narrativity consists of two key components: (1) what is conveyed and (2) how it is conveyed (Chatman 1978). Differences in these narrative elements can significantly impact transportation into the narrative and the resulting persuasion (van Laer et al. 2019). Narrative transportation leaves the audience with the perception that the content creator's story is based on direct experience, which typically makes novel information easier to understand and seemingly more intuitively truthful (Marsh and Fazio 2006). As a result, the audience tends to engage with the story less critically and is ultimately persuaded by it.

However, not all content in the creator economy is overtly narrative in form, yet "definitions of narratives and stories are highly fluid. The most traditional assumption is that a narrative consists of a series of connected events with a clear beginning, middle, and end. Generally, a narrative is considered a superordinate construct composed of a series of stories or events, which may be imaginary or factual. Whereas novels and short-form fiction are narratives, the factual accounts of scientific discoveries are likely less story-like but are still narratives nonetheless" (Boyd et al., 2020, p. 3). This perspective suggests that even fact-based content (e.g., New York Times articles, TED Talks, or U.S. Supreme Court opinions) can exhibit varying degrees of narrativity, albeit with different forms of staging, plot progression, and cognitive tension.

Relatedly, van Laer et al., (2019) emphasized that narrativity should be understood as a continuum rather than a binary distinction between story and non-story. Prior research has also distinguished between story-based and argument-based messages, while acknowledging that real-world content often blends both forms. Notably, Green and Brock (2000) found that whether the message was framed as fiction or fact induced neither narrative transportation nor belief change, highlighting that how a message is conveyed (i.e., narrative discourse) and processed matters more than whether it is explicitly fictional. For example, Krause-Galoni and Rucker (2024) showed that embedding factual arguments in a narrative structure can increase elaboration and persuasion without diminishing transportation. Similarly, prior research has shown that narrativity can be present in factual content such as online reviews (van Laer et al. 2019), advertisements (Escalas 2004a), and documentaries (Green, 2008).

EMOTIONALITY IN CONTENT

Prior research has advised content creators to craft emotional content to increase audience engagement (Berger et al. 2021, 2023; Lee 2021; Lin et al. 2021; Tellis et al. 2019). For example, broadcasters should smile more during their live broadcasts to enhance engagement (Lin et al. 2021). Audiences engage not only by liking, commenting, and sharing the content, but also by discussing it in subsequent conversations (Berger and Milkman 2012). Further, audiences produce emotional content via narrative reviews of their consumption experiences (Berger 2014; Melumad et al. 2019; Yin et al. 2017).

Literature on content emotionality can be organized along two themes. The first theme relates to determinants of the *level* of emotionality (low vs. high; e.g., Melumad et al. 2019; Mosley et al. 2024), and the impacts of this level on audience engagement (e.g., Lin et al. 2021) and downstream marketing outcomes (e.g., Lee 2021; van Laer et al. 2019). For example, people generate shorter content when using their phones (vs. computers) and therefore use less emotional language (Melumad et al. 2019). Content broadcasters who express happiness emotions induce happiness among the audience. Thus, broadcasters can increase audience engagement with more smiles (Lin et al. 2021).

The second theme focuses on the *type* of emotionality that boosts engagement (e.g., Berger et al. 2023; Berger and Milkman 2012). For example, high-arousal emotions (e.g., awe) make content more viral due to physiological arousal (Berger and Milkman 2012). Anxious, exciting, and hopeful language captures attention, whereas sad language discourages it, with uncertainty and arousal playing the mediating role (Berger et al. 2023).

The preceding discussion suggests that academics have paid less attention to whether *changes* in emotionality, rather than its valence and extremity, can affect engagement. Emotionality changes are paramount to creators' ability to engage their audience with narrative content, which forms the creator economy's foundation. Thus, this lack of academic attention is consequential in advising creators of what content to create to engage the audience. Our research addresses this lack of attention.

Emotionality Flips

We extend prior research by reasoning that audience engagement with *long* narrative content (e.g., podcasts and TED talks) is driven more by *emotionality flips*, the number of times emotionality flips direction (from a positive slope to a negative slope or vice versa) from one unit of content (e.g., a sentence) to another. These flips may reflect dramatic plot twists or significant character developments. The extended length and complexity of long narratives allow characters and plots to unfold gradually (Feiereisen et al. 2021), creating multiple points of emotional engagement. Additionally, the average emotionality of a long narrative tends to regress toward the mean (Morin and Acerbi 2017) because highly emotional sentences are balanced out by less emotional ones, making the average less insightful. These emotionality flips generate dynamic shifts in the audience's emotions, creating memorable, impactful moments that sustain interest throughout the story. These emotional highs and lows prevent monotony, offering moments of relief and reflection that enhance overall enjoyment.

Indeed, neuroscience research has indicated that sensitivity to stimulus gradients, such as changes in the cardiac cycle, affects the detection of emotional nuances and emotional intensity. Importantly, this sensitivity operates independently of stimulus intensity, underscoring the critical role of bodily states and their variations (Garfinkel et al. 2014). Furthermore, experimental research has shown that changes in stimulus intensity affect cortical activity, the brain's process for handling information, making decisions, and controlling actions. When experiment participants rated stimuli with varying intensity, the variations in intensity—and not the average intensity—influenced perception and neural activity (Torta et al. 2020). Moreover, other studies have shown that changing the stimuli results in better brain responses to sensory input, indicating that variations in the stimulus intensity are a more effective measure than the average intensity. For instance, brains distinguish changes in sounds independently of their volume levels, demonstrating that variation in change is a superior metric (Whiteford and Oxenham 2023). Another study highlighted that the brain's sensitivity to changes in sensory differences, rather than absolute intensity, is a widespread feature of neural processing (Somervail et al. 2021).

Previous research has suggested that sentiment volatility enhances audience engagement with content (i.e., virality; Berger et al., 2021), indicating that sentiment volatility (i.e., the standard deviation of sentiment differences between adjacent chunks of content) can increase virality. Consistent with the notion that sentiment volatility makes experiences more stimulating, thereby increasing virality, we reason that not only the magnitude of emotion and its extremity (i.e., two components of volatility) but also the *changes in the direction of emotionality* impact audience engagement.

Our Measure of Emotionality Flips

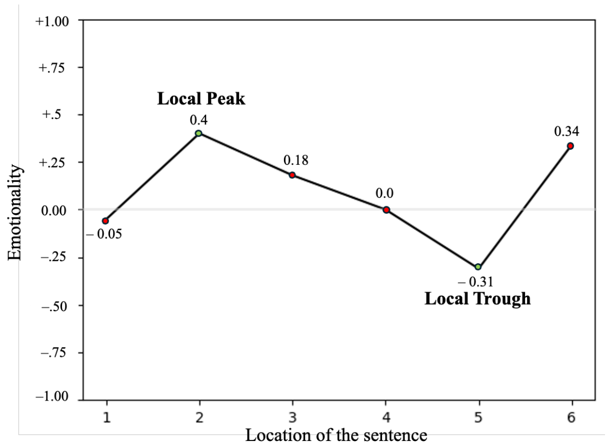
Three reasons drive us to measure emotionality at the *sentence*—as opposed to phrase or paragraph—level (Sepehri et al. 2021; van Laer et al. 2019). First, sentences represent natural, coherent units of thought and emotion within a text (Kleres, 2011). Theoretically, sentences encapsulate complete ideas and their corresponding emotionality, facilitating a more precise measurement of changes in emotionality valence (Bestgen 1994). This "completeness" is crucial for identifying temporal changes in emotions that drive audience engagement in long narratives. Second, sentences are typically neither too small (like a phrase), risking the loss of context and meaning, nor too large (like a paragraph), and may thus reveal the nuances of emotionality changes (Gee and Grosjean 1984). Therefore, a sentence is an ideal unit for analyzing the complexity and development of emotional arcs in narratives, offering a more accurate and detailed understanding of how emotionality impacts audience engagement. Third, the natural boundaries of sentences align well with how the brain processes and responds to language, making them an adequate chunk size for studying neural and psychological responses to narrative content (Ptaszynski et al. 2013).

Consider the following example of a six-sentence segment of a podcast. The example highlights the narrativity imperative in content.

> *"What happens if you leave school? You're so excited, school's over. You're walking home, just chatting with all your friends, and all of a sudden, a complete psychopath approaches you and pulls a gun on you. How would you react? What would you do in that moment, wondering what events are about to unfold from this psychopath on the loose? That is the case that I am excited to discuss today."*

Following previous research (Lacka et al. 2022; Pancer et al. 2019), we measure each sentence's emotionality score using the Valence Aware Dictionary and Sentiment Reasoner's (VADER's) *compound* variable because VADER measures the valence of words (i.e., positive, negative, and neutral), as well as their intensity (e.g., "okay" is rated 0.9, "good" 1.9, and "great" 3.1; Hutto and Gilbert 2014). Our code produces an array of six scores, one for each sentence: [−.05, .40, .18, .0, −.31, .34]. Figure 1 shows each sentence's location in the narrative on the X-axis and the sentence-specific emotionality on the Y-axis.

Figure 1. An example of emotionality flips



We are interested in counting the number of times emotionality changes direction. The first sentence has a slightly negative emotionality score of −.05. In contrast, the second sentence has a positive score of .40. Theoretically, as the audience moves from the first to the second sentence, they experience a positive shift in their emotional state. However, because the audience has heard only two sentences thus far, they have experienced no change (because "change" requires at least three points). The third sentence's score of .18 is less positive than the second sentence. In other words, emotionality has changed direction. The graph was upward sloping from the first sentence to the second. The second sentence has diminished the audience's emotional experience. This change in the direction of the emotionality creates a "local peak" in the graph. Our count of the number of times emotionality changes direction is now 1.

The emotionality score of the fourth sentence, which is lower than the emotionality score of the third sentence (0.0 < 0.18), indicates a consistent direction in the emotionality between these two sentences and, thus, no local optimum in the emotionality graph. The fifth sentence expresses the negative emotionality of −.31. In contrast, the sixth sentence conveys a positive emotionality of .34, creating a "local trough." That is, the second—and the last—

change in emotionality occurs when the audience moves from the fifth sentence to the sixth. Thus, our final count of the number of times emotionality flips direction is 2.

Next, we reason that what engages the audience is not the mere count of the times emotionality flips direction, but the *intensity* of this count throughout the narrative (Van Boven et al. 2010). Therefore, we divide the number of local optima by the number of sentences; that is, 2 ÷ 6 = 0.33. The reason for using intensity is that an episode with many sentences might naturally have more emotionality flips simply because of its length.[4] Consider two episodes: Episode A has 100 sentences and 20 emotionality flips, whereas Episode B has 200 sentences and 30 emotionality flips. Episode B has a greater number of changes in emotionality direction than Episode A (30 vs. 20). However, Episode A has 0.2 intensity of emotionality flips (20 ÷ 100), whereas Episode B has 0.15 intensity (30 ÷ 200). Therefore, we follow previous research that used proportions instead of counts of their textual variables (Packard et al. 2023; van Laer et al. 2019).

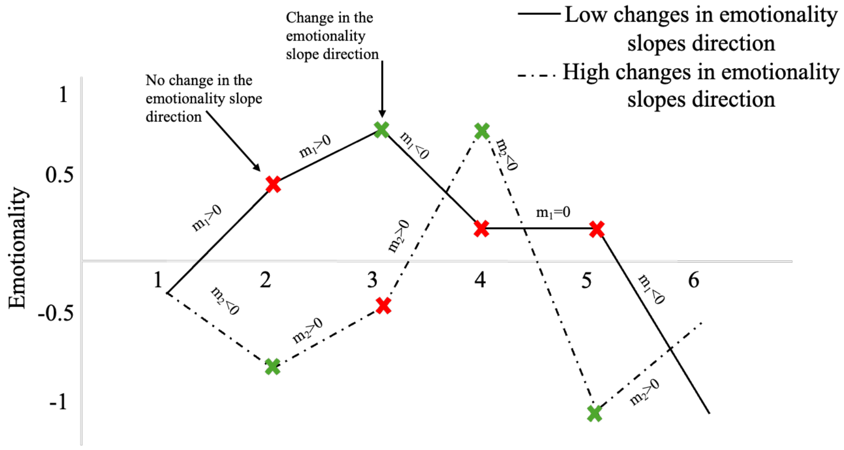
The preceding example illustrates two points that highlight the novelty of emotionality flips, relative to emotionality valence and intensity (Melumad et al. 2019; van Laer et al. 2019). First, the valence shifts three times, but emotionality changes direction twice. Therefore, our measure differs from merely counting the number of times the emotionality *valence* flips. We believe that measuring changes in the emotionality slope is a better metric than valence flips because the emotionality of content can remain positive (negative) even when the previous sentence had a positive (negative) score. Yet, the audience senses a downward (upward) slope, despite the valence being the same. For instance, after the sentence “After getting to know him better,” with an emotionality score of 0.44, one might say, “I liked him,” with an emotionality score of 0.42, or “I loved him,” with an emotionality

[4] Duration and length (number of sentences or words) are different variables that could be associated with engagement but might also influence the number of emotionality flips. Therefore, it is important to control for both the number of sentences and the duration when measuring emotionality flips, even when the independent variable is the intensity of emotionality flips.

score of 0.59. Depending on the content creator's choice of words, the slope could shift, but in both cases, the valence remains positive. Second, the emotionality intensity does not impact our measure, as even subtle changes in the emotionality slope direction can create an emotionality flip. This flip can be considered the highest (lowest) point of emotionality that an audience feels over the last sentence(s), after which they experience a downward (upward) trend. More concretely, the intensity of change between the fifth and sixth sentences is greater than between the third and second. However, these two sentences contribute to only a local extreme.

Mathematically, the changes in emotionality direction create a local optimum ($y'_n \neq y'_{n+1} \neq 0$).[5] A content with n sentences will produce $n-1$ slopes and $n-2$ positions for the local optimum (see Figure 2). In this way, each time there is an emotionality local optimum, we count it as a sudden emotional change. Last, we divide the number of local optima by the number of sentences to account for the frequency and distribution of optima, thereby more accurately measuring how dynamic and stimulating the content is.

Figure 2. Example of changes in the direction of emotionality slopes



Note. Both examples change the emotionality valence twice and have the same peaks.

Arousal and surprise as psychological mechanisms

Building on dual-process theories of information processing (e.g., Pham 1998;

[5] If there is no change in emotionality between two sentences, the derivative of the slope will be zero, which does not account for a local optimum in the emotionality graph.

Schwarz 2000), we distinguish between affective and cognitive mechanisms through which emotionality flips enhance audience engagement. We conceptualize arousal as an affective experience that reflects the visceral intensity of emotion (Mandler 1982), operating quickly and automatically to heighten engagement (Berger 2011; Pham 1998). In contrast, surprise represents both an emotional and cognitive response (Teigen and Keren 2003; Valenzuela et al. 2010)—an emotion triggered by expectancy violations that simultaneously initiates cognitive appraisal and mental updating (Fisk 2002).

Arousal is a condition in which one is physiologically alert, conscious, and attentive (Heilman 1997). This increase in alertness captures one's subjective feelings, thereby enhancing engagement (Schachter and Singer 1962). Additionally, arousal enhances attention, encourages content consumption, and increases the virality of content (Berger et al. 2023; Berger and Milkman 2012). For example, readers who encounter more arousal cues in a review are likely to infer that the reviewer experienced higher arousal while writing it, leading them to form related opinions about the reviewer and the review (Yin et al. 2017). Applied to narrative content, emotionality flips may serve as affective signals of arousal, indicating that the content creator experienced an emotional high or low, thereby transmitting emotional intensity to the audience and amplifying their engagement.

Surprise, in turn, is an emotion (Ekman et al. 1972) that arises from unexpected events, engaging cognitive processes of comparison and updating (Teigen and Keren 2003; Valenzuela et al. 2010). It reflects a hybrid mechanism—an emotional trigger with cognitive consequences—that captures how audiences process unexpected emotional turns. This surprise can enhance content creators' effectiveness by leveraging the content's novelty (Kim et al. 2024). Listeners anticipate (1) an increase in the positivity of the emotionality as the content progresses with a positive emotional trajectory or (2) a decrease in emotion when the content's emotion is on a downward trajectory. However, when the trajectory unexpectedly shifts, listeners experience surprise. For example, the audience may be surprised when listening to the Titanic story and suddenly witness the disaster unfold as the ship hits an iceberg while passengers are engaged in regular activities. Another example is the Red Wedding scene in the Game of Thrones series, where disaster strikes unexpectedly during a wedding, making it one of the most shocking scenes in the entertainment industry (Hough 2019). These unexpected emotional turns disrupt listeners' predictions, triggering cognitive appraisal processes (Teigen and Keren 2003) and increasing engagement by drawing renewed attention and interest.

In summary, emotionality flips increase engagement through (1) fast, automatic affective responses (arousal) and (2) slower emotionally triggered cognitive appraisals (surprise), preventing habituation and keeping the audience mentally invested in the content (Figure 3).

Fluency as a moderator

Individuals process narratives as if constructing and visualizing a story (Nielsen and Escalas 2010), naturally imposing temporal structure and causal links on events (Escalas 2004b). Therefore, we hypothesize that the narrative's fluency—the ease with which the audience processes content (Schuster et al. 2016)—moderates emotionality flips' effect on audience engagement.

We reason that emotionality flips unfold at the sentence level, making sentence-based fluency directly relevant to the audience's processing of emotional shifts. Prior research has demonstrated that processing fluency induces arousal through the pleasure of effortless comprehension (Forster et al. 2016). Critically, building on dual-process theories (Pham 1998; Schwarz 2000), highly fluent narrative reduces the cognitive resources needed for understanding (Alter and Oppenheimer 2009), thereby freeing up mental capacity for affective (i.e., visceral) responses to emotionality flips (van Laer et al. 2014).

In contrast, less fluent narrative taxes the audience's cognitive resources, reducing their ability to process and react to emotional shifts. Put differently, when content is fluent, reduced processing load enhances the affective resonance of emotionality changes while also increasing general audience liking (Figure 3).

Need for Affect as a moderator

Individuals differ in their motivation to approach or avoid emotionally charged experiences, a trait captured by *Need for Affect* (NFA; Maio and Esses 2001). High-NFA individuals are motivated to seek out, experience, and make sense of emotions, whereas low-NFA individuals tend to avoid emotional situations, preferring affect-neutral experiences. These motivational orientations shape how people process affective information and respond to emotionally laden narratives (Appel et al. 2012).

Evidence from behavioral and neurocognitive research underscores that NFA determines the depth and style of affective processing. In naturalistic reading paradigms, high-NFA individuals exhibit slower reading of emotion-laden cues, reflecting deeper engagement and elaboration on affective cues, whereas low-NFA individuals process such cues more rapidly (Lei et al. 2023). Similarly, high-NFA individuals show greater behavioral responsiveness to affective variable messages than their low-NFA counterparts (Haddock and Maio 2019).

Extending these findings to narrative dynamics, we reason that emotionality flips serve as affectively rich signals that evoke arousal and surprise. Such affective variability is inherently engaging for high-NFA individuals because it aligns with their motivation to approach, experience, and interpret emotions. By contrast, low-NFA individuals are less inclined to seek out affective stimulation and may disengage when confronted with emotional shifts, perceiving them as overwhelming. Thus, NFA's motivational role determines whether emotionality flips enhance or dampen audience engagement with the content (Figure 3).

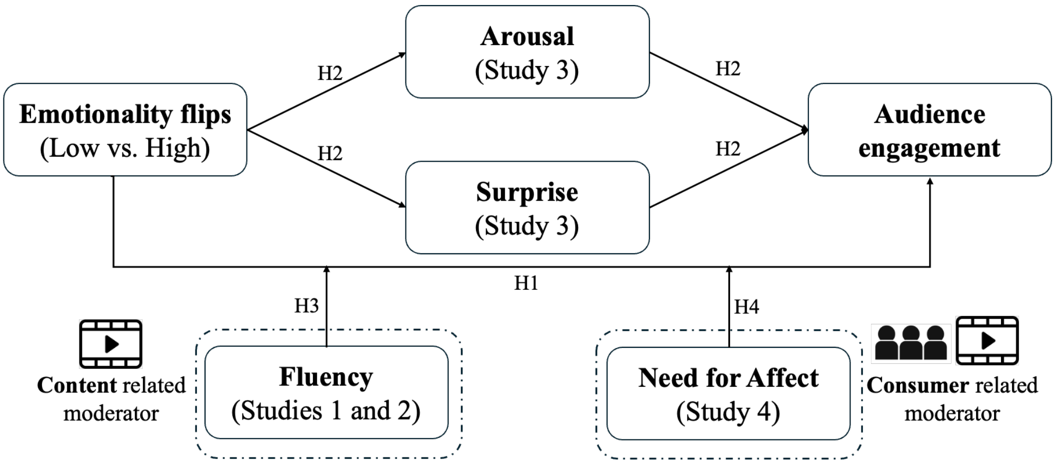
H_1: *Emotionality flips increase audience engagement with long narratives.*

H_2: *Audience arousal and surprise mediate emotionality flips' effect on audience engagement. In other words, emotionality flips increase audience arousal and surprise, each of which in turn increases audience engagement.*

H_3: *Narrative fluency strengthens emotionality flips' positive effect on audience engagement. In other words, emotionality flips' positive effect on audience engagement is stronger when fluency is high (vs. low).*

H_4: *Need for affect moderates emotionality flips' positive effect on audience engagement, such that the effect disappears for people with a low (vs. high) need-for-affect.*

Figure 3. Conceptual framework



STUDY 1: PODCASTS ON YOUTUBE

Study 1 uses podcasts as the content of interest. A podcast is a digital medium that features spoken content on various topics, often serialized in formats such as interviews, discussions, or storytelling (Cox et al. 2023). A podcast is distinctive in its blend of narrativity and community engagement (Cox et al. 2023). In 2019, 274.8 million people listened to podcasts. By 2024, the number had grown to 504.9 million, an increase of more than 80%, and the market size reached US$30 billion. Furthermore, it is projected to exceed $130 billion by 2030 (Backlinko Team 2024). These statistics suggest the economic significance of research on podcasts.

Data and Method

We obtained data from YouTube on April 29, 2024. Specifically, we selected the 50 most popular English text-based podcasts as of April 29, 2024 (see Web Appendix B; Table

B1 lists names of the 50 podcast channels in our sample, along with the number of episodes in each channel). We collected the transcripts of each podcast's episodes published from January 1, 2019, to April 29, 2024.[6,7] This procedure yielded transcripts for 33,598 episodes. We measure the independent variable, the moderator, and the dependent variable at the episode level. Thus, the podcast *episode* serves as our unit of analysis.

Dependent Variable (DV) of *Audience Engagement*: Following previous research (Rizzo et al. 2023a), we measure audience engagement by the sum of the episode's (1) number of views, (2) the number of likes, and (3) the number of comments as of April 29, 2024. The sum is highly skewed and has high kurtosis (details in Web Appendix B, Table B2). Therefore, following prior research (Chen and Berger 2013), we natural log-transform (i.e., log transform with base *e*, where *e* is 2.718) the values of the sum (see Web Appendix B; Figure B1 presents the histogram of the values of audience engagement before and after the log transformation). Web Appendix C's Table C1 reports robustness across the three individual counts (i.e., logarithm of the number of views, likes, and comments) as DVs.

Explanatory Variable of *Emotionality Flips*: We measure an episode's emotionality flips by a ratio variable. The numerator is the number of times sentence-level emotionality changes direction in the episode. The denominator is the episode's number of sentences.

We calculate the values in six steps. First, we split each episode into sentences (Packard et al., 2023). Second, prior marketing research has reasoned that "VADER outperforms other commonly used benchmark methodologies such as Linguistic Inquiry and Word Count, Affective Norms for English Words, and the machine learning algorithm support vector machine" (Lacka et al., 2022, p. 69). Therefore, we measure each sentence's emotionality using the Valence Aware Dictionary and Sentiment Reasoner's (VADER's)

[6] We omitted podcasts that do not include text (e.g., music podcasts).
[7] Podcasts have experienced a rise in popularity since 2019 (Cramer-Flood 2021).

compound variable (Hutto and Gilbert 2014). VADER classifies words by valence (i.e., positive, negative, and neutral) and weights the valence by intensity (Pancer et al. 2019).[8] Third, we draw a graph, plotting sentence number (e.g., 1, 2, and so on) on the X-axis and the sentences' *compound* values on the Y-axis. Thus, each sentence occupies a point in the two-dimensional space. We connect the points to form a zigzag line representing the episode's emotional trajectory. Fourth, we calculate the slope between each pair of points. Fifth, we compute the number of times the slope changes from upward to downward (i.e., from positive to negative) or vice versa. Sixth, we divide the number of times emotionality flips (i.e., changes direction) by the number of sentences.

Moderator Variable of *Fluency*: We measure a narrative's fluency by the Flesch Reading Ease metric (Flesch 1948), a widely adopted measure of processing fluency (Markowitz 2023; Packard et al. 2023; Pancer et al. 2019). The Flesch Reading Ease score is calculated based on (1) the average sentence length and (2) the average number of syllables per word, with higher scores indicating more fluent (i.e., easier-to-process) text (Markowitz 2023; Schuster et al. 2016; see Web Appendix D for the formula).

Control variables. We list our control variables in four groups: (1) 25 episode-specific control variables related to the episode's *topics*, (2) five" other" episode-specific covariates, (3) one podcast-specific control variable, and (4) four levels of fixed effects (channel, year, month, and weekdays). Table 1 lists the variables (see Web Appendix B; Table B4 for descriptive statistics and Table B5 for Pearson bivariate correlation coefficients).

Table 1. Study 1: Variables

Variable	Measure	Reason for inclusion, if regressor	Reference research
DV			
Audience Engagement$_e$	The sum of three counts: (1) the episode's number of views, the number of likes, and the number of		Rizzo et al. (2023a)

[8] For example, the compound score for "love" is 0.63, whereas the compound score for "hate" is −0.57. Similarly, the compound score for "excellent" is 0.57, but the compound score for "terrible" is −0.47.

	comments as of April 29, 2024.		
Moderator			
Fluency$_e$	The ease with which *words* and *sentences* are processed (determined by their length and syntactic complexity).	Moderator	Markowitz (2023); Schuster et al. (2016)
Controls			
Average emotionality$_e$	The value of the VADER *compound* score, averaged over all sentences of the episode.	Overall emotionality of the content can be associated with engagement.	Berger et al. (2023); Berger and Milkman (2012)
Duration$_e$	Episode's number of seconds.	Previous research has suggested that content duration can be associated with engagement.	Markowitz and Shulman (2021)
Number of subscribers$_e$	Controlling for the number of subscribers to the channel.	The more subscribers the podcast channel has, the more the audience engages with the episodes.	
25 Topic weights$_e$	25 episode-specific control variables related to the episode's topics (see Web Appendix E for details).	Audience engagement with content is a function of the topics it covers.	Berger et al. (2021); Packard et al. (2023)
Channels FEs	Channel ID	It helps us control for time-invariant, channel-specific determinants (e.g., topics, podcaster) of audience engagement with an episode.	Perse (1998)
Years FEs	2019 to 2024	Audience engagement may vary by year (e.g., during the COVID-19 pandemic).	Li et al. (2021)
Months FEs	12 months of the year	Audience engagement may vary by the month (e.g., weather)	Stroud and Muddiman (2019)
Weekdays FEs	7 days of the week	Audiences may engage less on their workdays than on their off-work days.	Devereux et al. (2020)

Note: Subscript e refers to episode.

Regression Specification

To estimate the effect of emotionality flips on audience engagement, we include episode-level covariates and fixed effects for channels, years, months, and weekdays. Because emotionality flips is non-normally distributed and may be endogenous to audience engagement, we include a Gaussian copula term in the specification (Park and Gupta 2012; details of Normality tests are provided in Web Appendix B, Table B3, and Figure B2). Therefore, we specify the following equation to measure a podcast episode e's emotionality flips' effect on audience engagement with the episode e, which was published on YouTube channel c on day d of month m in calendar year y.

$$Log\ (Audience\ engagement) = \beta_0 + \beta_1 \times (Emotionality\ flips)_e + \beta_2 \times (Fluency)_e + \beta_3 \times (Number\ of\ sentences)_e + \beta_4 \times (Average\ emotionality)_e + \beta_5 \times (Number\ of\ days\ since\ posted)_e + \beta_6 \times (Duration)_e + \beta_7 \times (Number\ of\ subscribers) + \beta_{8-32} \times (Topic\ weights)_e + \beta_{33} \times (Copula\ term\ for\ the\ emotionality\ flips)_e + Channel_c + Day_d + Month_m + Year_y + \epsilon_e \tag{1}$$

The episodes within the same podcast are nested within the channels and thus not independent of each other. Therefore, we estimate standard errors clustered at the channel level. Lastly, our DV is a log-transformed count. Consequently, we use the ordinary least squares (OLS) estimator to estimate Equation 1.

Results

Main specification. Table 2's Column I reports the estimates. We find a positive association between the emotionality flips in a podcast episode and audience engagement with the episode (Table 2, Column I: $b_{Emotionality\ flips \rightarrow Audience\ engagement} = 0.945$, $p < .001$, Exp(b) = 2.57). The interpretation is that a 1% increase in emotionality flips in a podcast episode is associated with a 157% increase in engagement with the episode. This result supports H_1. The estimated coefficient of the copula term is insignificant, suggesting that *Emotionality flips* is likely exogenous in our specification (Papies et al. 2017).

Moderation by Fluency. The estimated coefficient of the interaction term between emotionality flips and fluency is positive and statistically significant, providing support for H_3 (Table 2, Column II: $b_{Emotionality\ flips \times Fluency \rightarrow Audience\ engagement} = 0.0001$, $p < .01$). This finding supports our theorization that fluency facilitates emotionality shifts' affective processing by reducing cognitive load and enhancing the audience's capacity to engage with emotionally dynamic content.

Table 2. Study 1: Regression Estimates

	DV = log (Audience engagement)	
	(I) Main specification	(II) Moderation
Emotionality flips	0.945*** (0.341)	1.028*** (0.339)
Fluency	−0.00004*** (0.0000)	−0.0002* (0.0001)
Emotionality flips × Fluency		0.0001** (0.0001)
Number of sentences	0.0004** (0.00003)	0.0004** (0.0001)
Average emotionality	−0.661*** (0.133)	−0.659*** (0.132)
Days elapsed	0.002*(0.001)	0.002* (0.001)
Duration	0.0001*** (0.0000)	0.0001*** (0.0002)
Number of subscribers	−0.0000 (0.0000)	−0.0000 (0.0000)
Copula term for IV	−0.003 (0.010)	−0.010 (0.010)
25 LDA Topic weights	✓	✓
Channel FEs	✓	✓
Year FEs	✓	✓
Month FEs	✓	✓
Day FEs	✓	✓
Constant	7.861*** (1.828)	7.826*** (1.828)
Observations	33,584	33,584
Log likelihood	−52,677.720	−52,674.090
AIC	105,563.400	105,558.200
Note:		$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Robustness Checks

We conducted an extensive set of robustness analyses to ensure that the relationship between emotionality flips and audience engagement is not an artifact of model specification, measurement choices, or alternative linguistic constructs. Across nine robustness checks—including (1) random-effects specification with channel treated as a random effect, (2) an alternative rule-based sentiment tool (TextBlob), (3) a deep-learning–based sentiment regressor (DistilBERT), (4) ChatGPT-based emotionality scoring, (5) alternative units of emotionality measurement (two-sentence chunks),[9] (6) alternative dependent-variable scaling, (7) controls for sentiment volatility (Berger et al. 2021), (8) controls for the magnitude of emotionality changes, and (9) controls for semantic progression (Toubia et al.

[9] See Web Appendix F, Table F1 (columns X–XII) for results using three-, four-, and five-sentence chunk sizes. The results remained the same, demonstrating that the effect of emotionality flips on engagement is robust to alternative chunking specifications.

2021)—the positive association between emotionality flips and audience engagement remained statistically significant and directionally consistent (Web Appendix F; Table F1). The effect sizes obtained from these alternative approaches closely track those from our main specification, demonstrating that the observed effect is robust to multiple modeling assumptions, alternative sentiment engines, and theoretically relevant covariates.

Study 2: TED TALKS

Influencers are regarded as opinion leaders due to their extensive reach and ability to shape public perceptions by disseminating content across social media platforms (Leung et al., 2022b; Mardon et al. 2023). They aim to influence their followers' opinions and behaviors. This influence is directly tied to the level of engagement their content can generate. Therefore, our Study 2 samples influencers' content (in place of podcasts), testing whether influencers can boost audience engagement by flipping emotionality in their content and whether fluency can moderate this effect.

We source data from TED.com, a platform dedicated to the dissemination of "ideas worth spreading." Prior research has examined the relation between a TED talk's content and its number of views (Markowitz and Shulman 2021). Viewers are more likely to share an engaging talk with their online network (e.g., on Facebook or LinkedIn). Such sharing, in turn, increases the talk's number of views (Seiler et al. 2017).

Data and Method

We use a publicly available dataset of 3,381 TED talks.[10] Thus, a talk is our unit of observation and analysis. The measurements for the independent variable, moderator, and dependent variables in Study 2 are identical to those in Study 1, with the slight difference that we measure engagement by the sum of (1) the number of views and (2) the number of

[10] https://github.com/The-Gupta/TED-Scraper

comments. The sum is highly skewed and has high kurtosis. Therefore, we natural log-transform the values of audience engagement (our DV) and estimate the regression using the OLS estimator (see Web Appendix B; Table B2 for details, and Figure B3 presents the histograms of audience engagement before and after log-transformation). Web Appendix C's Table C2, reports robustness using (1) the log of the number of views and (2) the log of the number of comments as DVs.

We retain all covariates from Study 1 except (1) the number of subscribers and (2) channel-fixed effects, which do not apply to TED talks. We include two new control variables. First, we control for the number of available subtitles for the talk because the more subtitles available, the more accessible the talk is to audiences with varying language preferences. Second, we included fixed effects for speakers' gender. Furthermore, using the same topic modeling method as in Study 1, the 20-topic model yielded the lowest perplexity score (see Web Appendix B, Table B6 for descriptive statistics and Table B7 for correlation coefficients).

Our specification likely omits variables that correlate with emotionality flips and audience engagement. We account for *Emotionality flips*' potential endogeneity by using the Gaussian Copula method (Park and Gupta 2012; Web Appendix B, Table B3 details the normality tests). Each of the four tests rejects the null hypothesis of *Emotionality flips*' normality and thus supports our inclusion of the Gaussian copula term to control for *Emotionality flips*' potential endogeneity (see Web Appendix B; Figure B4).

Results

Regression estimates report that a Ted talk's number of emotionality flips is positively associated with audience engagement with the Talk (Table 3, Column I: $b_{Emotionality\ flips} \rightarrow$

$_{Audience\ engagement}$ = 0.828, $p < .001$, *Exp*(b) = 2.29).[11] The interpretation is that a 1% increase in the emotionality flips in a TED talk is associated with a 129% increase in engagement.[12] This result supports H_1. The estimated coefficient of the copula term is insignificant, suggesting that *Emotionality flips* is likely exogenous in our specification (Papies et al. 2017).

Moderation by Fluency. We next test whether the Talk's fluency strengthens emotionality flips' positive effect on audience engagement. The results, reported in Table 3, indicate a significantly positive coefficient of the interaction between emotionality flips and fluency (Table 3, Column II: $b_{Emotionality\ flips\ \times\ Fluency\ \rightarrow\ Audience\ engagement} = 0.006$, $p < .05$), providing support for H_3. This finding aligns with our theoretical argument that narrative fluency facilitates affective processing of emotional variation, enabling audiences to experience and respond to shifts in emotional tone more fully.

Table 3. Study 2: Regression Estimates

	DV = log (Audience engagement)	
	(I) Main specification	(II) Moderation
Emotionality flips	0.828*** (0.172)	0.504* (0.266)
Fluency	0.0002 (0.0003)	−0.0003 (0.0004)
Emotionality flips × Fluency		0.006* (0.003)
Number of sentences	0.003*** (0.0004)	0.0003*** (0.001)
Average emotionality	0.030* (0.017)	0.032* (0.017)
Days elapsed	0.003** (0.001)	0.003** (0.001)
Duration	0.0004*** (0.0001)	0.005*** (0.0001)
Number of subscribers	0.063*** (0.002)	0.063*** (0.002)
Copula term for IV	−0.005 (0.017)	−0.008 (0.017)
20 LDA Topic weights	✓	✓
Gender FEs	✓	✓
Year FEs	✓	✓
Month FEs	✓	✓
Day FEs	✓	✓
Constant	−6.439 (6.807)	−6.174 (6.809)
Observations	3,381	3,381
Log likelihood	−3,172.399	−3,170.842
AIC	6,468.799	6,467.684

[11] Results hold when the DV is the individual components of (1) the number of views, and (2) the number of comments, (see Web Appendix C; Table C2).

[12] $(Exp\ (b) - 1)\ \times\ 100 = (Exp\ (0.828) - 1)\ \times\ 100 = 129\%$.

Note:	$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Robustness Checks

We reported nine robustness checks for Study 1: (1) random-effects regression, (2) an alternative method (i.e., TextBlob) of measuring emotionality, (3) using a deep learning-based method to measure emotionality, (4) using ChatGPT-generated emotionality scores, (5) an alternative unit of measurement of emotionality (i.e., two-sentence instead of one-sentence chunks), (6) an alternative measure of the DV, (7) controlling for Berger et al.'s (2021) sentiment volatility, (8) controlling for the magnitude of emotionality flips, and (9) controlling for Toubia et al.'s (2021) semantic progression variables.

Unlike podcast episodes, which are nested within channels, TED Talks are not embedded within higher-order structures such as channels. Therefore, a random-effects regression is not particularly appropriate in this context. However, the remaining nine robustness checks are applicable. Results from these tests continue to support our main findings. Table F2 in Web Appendix F reports estimates consistent with those presented in Table 3, Column I of the main specification.[13]

Study 3: REPLICATING THE EFFECT IN A CONTROLLED EXPERIMENT

Study 3 tests the effect in a controlled lab setting, leveraging random assignment to examine the causal interpretation of our hypothesized coefficients in H_1. Moreover, it helps test the underlying mechanisms by examining both affective and cognitive pathways through which emotionality flips may influence engagement (our H_2). To rule out curiosity as a hybrid motivational state arising from perceived knowledge gaps (Berlyne 1966), we

[13] We also tested the measure using three-, four-, and five-sentence chunk sizes. The results remained consistent across all specifications, confirming the robustness of our findings. See Web Appendix F, Table F2 for detailed results.

measured curiosity alongside arousal and surprise to ensure that any observed effects of emotionality flips on engagement were not better explained by heightened curiosity.

Participants, Design, and Procedure

Our pre-registered experiment (see https://aspredicted.org/tzz7-5hz3.pdf) recruited 801 participants (M_{Age} = 42.13, 48.4% female) from Prolific and randomly assigned them to one of two between-subjects conditions: low emotionality-flips vs. high emotionality-flips.

We rely on previous experimental research on long-form content (e.g., TED Talks; Sepehri et al. 2025) to create stimuli for manipulating emotionality flips. Specifically, we utilized Study 1's emotionality flips scores and randomly selected two podcast episodes, one from the top 10% (emotionality flips' score = 0.66) and the other from the bottom 10% (emotionality flips' score = 0.36) of the distribution. We drew these stimuli-episodes from the *same* podcast channel to control for potential confounds related to topic, host style, and episode length, thereby ensuring that only the emotionality flips' level varies across the two experimental conditions (see Web Appendix G for stimuli). We assigned each participant randomly to a condition. The participant listened to a podcast episode and subsequently completed the survey.

While Studies 1 and 2 measured audience engagement by the sum of the number of views, likes, and comments received by each podcast/talk, Study 3 measures engagement by interest and liking (i.e., "How interesting did you find this podcast episode? How much did you like this podcast episode?" α = .97; Sepehri et al. 2025).

We also measured several psychological mechanisms hypothesized to mediate emotionality flips' effect on audience engagement. Specifically, we measure *Arousal* using three bipolar scales that assess the degree of energy, activation, and excitement elicited by the episode (α = .85; Berger 2011). We measure *Surprise* by three items asking the extent to which participants felt able to predict upcoming events, how unexpected they found the storyline, and how difficult it was to anticipate what would happen next (α = .81). Lastly, we measure *Curiosity* by asking participants to report their desire to continue listening to the series, learn more about the narrator, and gain further information about the main character (α = .89; see Web Appendix G for items). We concluded the study by collecting participants' basic demographic information.

Results

Audience engagement. Results reveal that emotionality flips increase participants' engagement with the episode. Specifically, participants in the high emotionality-flips condition (M_{High} = 5.24, SD = 1.78) engaged with the podcast episode more compared to participants in the low emotionality-flips condition (M_{Low} = 4.90, SD = 1.80; F(1, 799) = 7.34, p = .007, partial η^2 = .01).

Arousal. Participants in the high emotionality-flips condition (M_{High} = 5.01, SD = 1.50) reported feeling higher arousal while watching the podcast episode than those in the low emotionality-flips condition (M_{Low} = 4.52, SD = 1.70; F(1, 799) = 18.31, p < .001, partial η^2 = 022).

Surprise. Participants in the high emotionality-flips condition (M_{High} = 3.98, SD = 1.08) reported feeling a higher level of surprise while watching the podcast episode than those in the low emotionality-flips condition (M_{Low} = 3.83, SD = 1.16; F(1, 799) = 3.68, p = .050, partial η^2 = .005).

Curiosity. Participants in the high emotionality-flips condition (M_{High} = 4.61, SD = 2.02) and the low emotionality-flips condition reported *similar* levels of curiosity (M_{Low} = 4.59, SD = 2.02; F < 1, p = .882, partial η^2 = .000). Therefore, this variable cannot be a mediator.

Mediation. Next, we examine whether arousal and surprise mediate emotionality flips' effect on participants' engagement with the podcast. First, we conduct separate mediation analyses using PROCESS Model 4 (Hayes 2022; 5,000 bootstrapped samples), testing each mediator individually. Results reveal that arousal (indirect effect = 0.21, 95% CI = [0.10, 0.29]) and surprise (indirect effect = 0.03, 95% CI = [0.01, 0.05]) significantly mediate the effect of emotionality flips on podcast liking. This result supports H_2.

We assess the relative contribution of each mechanism variable by including them simultaneously in a parallel mediation model using PROCESS Model 4 (Hayes 2022; 5,000 bootstrapped samples). The indirect effect via arousal remains significant (indirect effect = 0.20, 95% CI = [0.11, 0.29]), whereas the indirect effect via surprise is no longer significant (indirect effect = 0.00, 95% CI = [−0.00, 0.01]). These findings suggest that, while surprise may help explain the effects of emotionality flips, arousal is the primary mechanism underlying them. This result brings clarification to H_2.

Discussion

Study 3 replicates emotionality flips' causal impact on audience engagement in a carefully controlled lab experiment and identifies arousal as the key explanatory mechanism. While both arousal and surprise mediate the effect when tested separately, only arousal remains significant when the two variables are modeled simultaneously, accounting for the majority of the variance in liking. These results provide strong causal support for H_1 and H_2, demonstrating that arousal is the primary driver of emotionality flips' effect on audience engagement. Lastly, this mechanism aligns with our findings regarding the moderating role of fluency (H_3): disfluent content allocates more resources to the cognitive system and diverts resources away from the affective route, which served as the main pathway through which emotionality flips increased engagement.

Study 4: MODERATING ROLE OF NEED FOR AFFECT

Study 4 pursues three main goals. First, it seeks to replicate the main effect of emotionality flips on audience engagement using podcast stimuli we created for this study

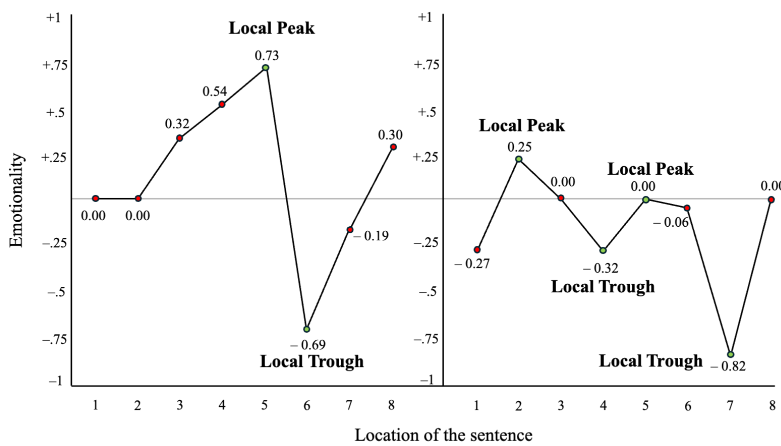
(H_1). Second, because these podcasts were generated using the same text-to-speech system and identical voices, the fluency, tone, and idiosyncratic vocal style of the podcaster were held constant, thereby ruling out these potential confounds in Study 3. Third, the study aimed to test the moderating role of NFA to examine whether the positive impact of emotionality flips on audience engagement depends on listeners' motivation to approach and experience emotions (H_4).

Participants, Design, and Procedure

Our pre-registered experiment (see https://aspredicted.org/sg7ft4.pdf) recruited 200 participants (M_{Age} = 45.42, 55.1% female) from Prolific and randomly assigned them to one of two between-subjects conditions: low emotionality flips vs. high emotionality flips.

To ensure that potential confounds such as vocal fluency, tone, or idiosyncratic voice characteristics were held constant, we created both podcast episodes ourselves using text-to-speech synthesis. This approach enabled us to manipulate only the degree of emotionality flips, while maintaining identical pacing, prosody, and vocal quality across conditions. Both podcasts narrated the same storyline but were systematically modified to differ in their emotional variability (excerpts in Web Appendix H). The two versions were identical in length, pacing, and voice but differed in the emotionality flips (Figure 4).

Figure 4. Emotionality flips across both conditions (Left: low emotionality flips; Right: high emotionality flips)



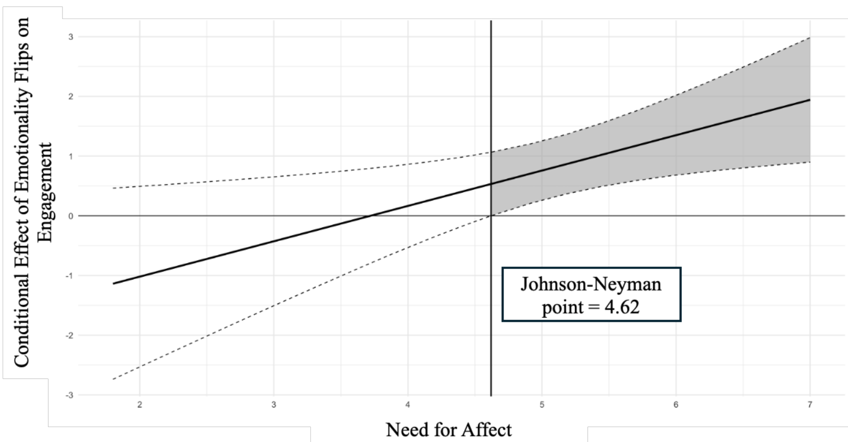
After listening to one of the two podcast versions, participants evaluated the episode using the same measures used in Study 3 (α = .97). Participants also completed the NFA scale (Appel et al. 2012), consisting of ten items (e.g., "I feel that I need to experience strong emotions regularly"; "I would prefer not to experience either the lows or highs of emotion" [reverse-coded]; full scale in Web Appendix H, α = .70). The study concluded with collecting basic demographic information.

Results

Audience engagement. Participants in the high-flip condition (M_{High} = 4.91, SD = 1.51) reported greater engagement than those in the low-flip condition (M_{Low} = 4.14, SD = 2.04; $F(1,196) = 9.22$, $p = .003$, partial $\eta^2 = .045$).

Moderation. To test the moderating role of NFA, we conducted a moderation analysis (PROCESS Model 1, 5000 bootstrapped samples; Hayes 2022). We find a significant interaction between emotionality flips and NFA (b = 0.59, SE = 0.24, t = 2.49, $p = .014$), indicating that the effect of emotionality flips disappears for people with low NFA. Floodlight analyses (Spiller et al. 2013) indicate that this effect holds only among participants with an NFA of 4.63 or higher (63.64% of the sample; Figure 5).

Figure 5. The conditional effect of emotionality flips on DV at different levels of Need for Affect. The grey area indicates the Johnson-Neyman region of significance



Discussion

Study 4 replicates the main effect of emotionality flips on audience engagement and provides evidence that this relationship is moderated by *Need for Affect*. Specifically, individuals high in NFA were significantly more engaged by podcasts featuring high emotionality flips, whereas the effect disappeared among individuals low in NFA. This pattern supports H_4. Additionally, by holding constant the fluency, tone, and idiosyncratic vocal style of the podcaster through text-to-speech synthesis, Study 4 rules out these potential confounds and demonstrates that the observed effects are driven by emotional dynamics rather than differences in delivery style or expressiveness.

Study 5: VALIDATION WITH SYNTHETIC AUDIENCES

With the rapid rise of generative AI, researchers increasingly use Large Language Models (LLMs) to simulate human judgments and generate synthetic behavioral data (Huang and Rust 2025). These models can approximate consumer evaluations with high reliability, offering a scalable, replicable complement to traditional data collection (Toubia et al. 2025). Building on this methodological advancement, Study 5 uses synthetic participants to examine whether podcasts with high emotionality flips are perceived as more engaging than those with low emotionality flips.

Design and Procedure

Study 5 employed a between-subjects design: low emotionality flips vs. high emotionality flips. We test the robustness of our findings using long-form stimuli by developing two podcast transcripts that describe the Apollo 11 Moon landing. The two versions were identical in length, topic, and informational content but differed systematically in emotionality flips: the low-flip version contained 20 emotional shifts across 65 sentences (flip score = 0.31), whereas the high-flip version included 43 emotional shifts across 62 sentences (flip score = 0.69; see Web Appendix I for full transcripts and Figure I1 for the sentence-level flip distributions).

We then used the ChatGPT API (model gpt-4o) to simulate 5,000 independent participants (2,500 per condition). Each synthetic participant was randomly assigned two or three human-like characteristics from a pool of 50 traits (e.g., curiosity, analytical thinking, emotional sensitivity, attentional control; full list in Web Appendix I, Table I1). These traits were embedded directly into the system prompt, instructing the model to respond as if it were a person possessing those characteristics (Yuan et al. 2024). After reading one version of the podcast transcript, each synthetic participant answered a single engagement question: "How engaging did you find this podcast episode?" (1 = Not at all engaging, 7 = Extremely engaging).

For each condition, the model generated 2,500 unique engagement ratings with randomized temperature settings (0.8 – 1.2) to induce natural variability across simulated respondents. Each response, along with the assigned characteristic codes, was recorded automatically via the API and stored in a structured dataset.

Results

Audience engagement. Results showed that synthetic participants exposed to the high-flip podcast rated it as significantly more engaging (M_{High} = 4.96, SD = 1.13) than those exposed to the low-flip version (M_{Low} = 4.87, SD = 1.15; F(1, 4,998) = 7.58, p = .006). The effect remained significant after controlling for all 50 personality-like traits (F(1, 2,449) = 32.05, p < .001).

Discussion

Study 5 validated the effect of emotionality flips using synthetic audiences simulated via ChatGPT. This approach demonstrates that the observed patterns generalize beyond traditional, limited samples and can be reproduced in large-scale, fully controlled, long-form narrative contexts.

GENERAL DISCUSSION

Our research develops the concept of emotionality flips. It tests its impact on audience engagement using two large-scale observational datasets, two controlled lab experiments, and a synthetic validation using LLMs. The five studies demonstrate that emotionality flips are associated with, and causally lead to, higher audience engagement, providing support for H_1. Study 3 also indicates that arousal drives this effect, thus supporting and clarifying H_2. Furthermore, our observational studies consistently reveal that narrative fluency moderates emotionality flips' effect, such that emotionality flips are more impactful when the content is easier to process, providing support for H_3. Lastly, Study 4 introduces a boundary condition based on individual differences, showing that the effect of emotionality flips depends on the audience's NFA (H_4), thereby extending the framework to account for individual variability in audience responses.

Contributions to Theory

First, previous research has documented various determinants of audience engagement with content. These determinants include (1) overall emotionality (Lee 2021), (2) specific types of emotionality (Mosley et al. 2024), (3) non-emotional cues such as selfies with a brand (Hartmann et al., 2021), and (4) the audience's level of closeness with the person posting the content (Dubois et al. 2016). However, these studies largely overlook the role of content length and narratives—a key consideration in the creator economy, where longer narratives often translate into greater monetization opportunities. Because revenue is closely tied to engagement duration, content creators are incentivized to produce long-form narratives that sustain audience interest. We contribute to this literature by introducing a novel construct that is especially relevant to long narratives: *emotionality flips*. Unlike valence, intensity, or emotional type, emotionality flips capture the dynamic flow of emotion over time, making them uniquely suited to understanding engagement in long-form content. By focusing on this temporal emotional structure, our work offers new, content-specific insights into how creators can craft more engaging narratives.

Second, our research contributes to the literature on narrativity by focusing on the discourse dimension (i.e., how narrativity is conveyed; Chatman 1978; Culler 1981)—specifically, the role of emotionality flips in narrative delivery. Emotionality flips serve as narrative discourse tools that introduce emotional dynamics into messages, allowing podcasters and speakers to embed a sense of narrativity and benefit from its persuasive effects (i.e., narrative persuasion; Argo et al., 2008; Escalas, 2007).

Third, we extend dual-process models (e.g., Pham 1998; Schwarz 2000) of engagement by uncovering the underlying psychological mechanisms through which emotionality flips operate. Study 3 reveals that arousal mediates the effect of emotionality flips on podcast liking, whereas surprise plays a weaker and less consistent role. This suggests that emotionality flips engage audiences primarily through fast, visceral emotional activation. This contributes to the theories of affective engagement (Berger 2011; Mandler 1982) by showing that dynamic emotional variation, when processed fluently, can trigger heightened arousal that sustains attention and liking.

Implications for Practice

We offer insights for content creators, influencers, opinion leaders, and anyone seeking to shape audience opinions and enhance engagement. Additionally, we cater to platforms that curate and recommend content, as well as to consumers who aim to select content that best fits their interests and goals.

Recommendations for content creators. Prior research has advised content creators to engage their audience by (1) defining a clear niche (Garcia 2021), (2) using specific types of words (e.g., semantic; Rizzo al., 2023a; Rizzo et al., 2023b), (3) delivering specific types of emotions (Lin et al. 2021), (4) referencing their close social ties (Chung et al. 2023), and (5) paying attention to the match between supported brands and followers (Leung et al., 2022a). Practitioners also offer recommendations, such as being creative (Adobe Express 2024), creating longer YouTube videos (Peterson 2024), and maintaining a strong connection with the audience (Spotify 2024).

However, academic and practitioner discussions have mostly overlooked the emotionality dynamics that unfold within a narrative. Our research fills this "gap" by demonstrating that emotionality flips—deliberate shifts between emotional highs and lows—increase audience engagement. Accordingly, we put forth the following recommendations:

1. *Creators should create emotional variation.* They should design content with emotional highs and lows and thus increase engagement.
2. *Creators should increase audience arousal.* Audience arousal is the primary underlying psychological mechanism. Therefore, creators can heighten arousal not only by flipping emotionality but also through nonverbal elements such as pauses, music transitions, vocal inflections, or visual pacing.
3. *Creators should create content that is easy to process for the audience.* Maintain narrative fluency by keeping transitions coherent and natural so that emotional shifts enhance, rather than disrupt, the audience experience.
4. *Creators should acknowledge individual differences in need-for-affect (NFA).* Emotionality flips are most effective for audiences with high NFA. Therefore, emotionality flips are beneficial for content targeted at high-NFA audiences, but not for content targeted at low-NFA audiences.

Recommendations for platforms. Our findings hold important implications for platform-level strategy and algorithmic design. Recommendation systems employed by platforms such as YouTube, Spotify, and TikTok often prioritize content that elicits high arousal or surprise (Berger et al. 2023; Teixeira et al. 2012). Emotionality flips can serve as a strong signal to such curation algorithms, increasing the likelihood that emotionally dynamic content surfaces on platform users' feeds. Algorithms may also tailor recommendations based on users' prior engagement with emotionally fluctuating narratives, meaning creators who strategically incorporate emotionality flips may benefit from greater algorithmic amplification. Accordingly, we put forth two recommendations for platforms:

1. *Platforms should incorporate the emotionality flips feature into recommendation algorithms*. Including emotionality flips as a predictive feature in audience engagement models can increase platform users' acceptance of the recommended content and thus boost the model's true positives. Additionally, content consumers often watch content to experience a sense of arousal. Therefore, including emotionality flips may increase users' adoption of the recommended content, thereby boosting their trust in the recommendations.
2. *Platforms should develop AI-powered diagnostic tools for creators.* Content creators may lack knowledge of emotionality flips in their content. Platforms can provide a tool that scores a content's emotionality flips score and identifies emotionally flat or overly volatile segments in text, audio, or video content. The content creator may revise the segments and publish the content that they believe has the emotionality flips they are comfortable with. Subsequently, the platform should provide evidence of the association between content's emotionality flips and audience engagement. The evidence will give feedback to the content creator and help them adjust the emotionality flips in their future content.

Recommendations for Consumers. Emotionality flips also carry practical implications for audiences who seek engaging, meaningful, and emotionally balanced experiences. Understanding how emotional variation shapes engagement can help audiences make more intentional choices about the content they consume. Accordingly, we recommend that consumers:

1. *Select emotionality flip content for high engagement.* The audience should choose creators or genres with high emotionality flips because they provide emotional arousal.
2. *Seek emotionality flip content if high on need-for-affect (NFA).* The audience should favor emotionality flip content, particularly if they experience high NFA.

Emotionality flips beyond the creator economy. Beyond influencer content, emotionality flips may serve as a versatile mechanism across broader consumer communication domains. In advertising, sequencing emotional highs and lows can sustain arousal and enhance persuasion (Berger and Milkman 2012). In online reviews, moderate emotional variation may enhance perceived effort and helpfulness (Yin et al. 2017), while excessive emotional variation may reduce trust. Similarly, on platforms like Yelp, where secondary metrics such as “love this” and “oh no” are tracked, emotionality flips may help reviewers shape reader perceptions and boost content visibility. Lastly, since high-arousal emotions are linked to greater WOM diffusion, emotionality flips may increase the likelihood of sharing, expanding the reach of emotionally dynamic narratives (Consiglio et al. 2018).

However, emotionality flips may not be universally effective. Poorly calibrated arousal fluctuations may overburden audiences, reduce trust in micro-influencers (Rizzo et al., 2023b), or provoke discomfort with novel innovations (Noseworthy et al. 2014). Thus, emotionality flips’ effectiveness may hinge on an optimal match between arousal dynamics, audience goals, and contextual factors.

Limitations and Future Research

Our research has several limitations that offer opportunities for future inquiry. Although we examined processing fluency and NFA as moderators, other factors—such as content type, platform features, audience characteristics, and cultural context—may also shape how emotionality flips influence engagement. Emotionality flips may operate differently across genres (e.g., educational vs. entertainment), platforms (e.g., YouTube vs. Spotify), or creator styles, and these structural differences may interact with emotional dynamics in meaningful ways.

First, as shown in Study 4, audience characteristics are particularly important. Arousal responses depend on expectations shaped by familiarity with creators, prior engagement, and individual preferences (Teixeira et al. 2012). Loyal followers may interpret emotional shifts as intentional, whereas audiences expecting consistency may find them jarring. Individual differences in arousal sensitivity (Blascovich 1990) further suggest that some users seek emotionally dynamic content while others avoid it. Future research should examine how such audience-level factors condition responses to emotional structure in narrative content.

Second, cultural values may influence whether emotional variation is interpreted positively or negatively. For instance, Western audiences often view emotional dynamism as engaging, whereas East Asian audiences tend to prefer emotional restraint (Valenzuela et al. 2010). Investigating cross-cultural differences in emotionality flip processing could yield meaningful insights for global content strategies.

Third, past work shows that older adults rely more on emotionally congruent cues (Noh and Isaacowitz 2013) and interpret ambiguity more positively (Shuster et al. 2017). Gender differences may also matter (Kovacheva et al. 2022). Our observational datasets lack demographic information, limiting our ability to test these moderators directly. Study 3 provided an initial test and found no significant moderation by age or gender, though exploratory floodlight analyses suggested effects may emerge among older viewers. Future research should explore demographic heterogeneity using datasets with richer individual-level information.

Fourth, emotionality flips may not always be beneficial. If content repeatedly generates emotional shifts, it may exhaust audiences. Future work should examine whether varying the frequency or intensity of flips affects fatigue, enjoyment, and satisfaction, and identify optimal emotional dynamics for sustained engagement.

Fifth, prior research differentiates between mega- and micro-influencers (Rizzo et al. 2023b). Our data draws from highly popular YouTube podcasts and TED Talks, which may bias effects toward well-established creators. Future studies should examine whether emotionality flips operate similarly for micro-influencers or creators at earlier stages of audience growth.

Sixth, YouTube podcasts and TED Talks include visual cues. Although our hypotheses focus on textual emotionality (Berger et al. 2021; Markowitz and Shulman 2021), emotional fluctuations can also be conveyed through tone, facial expression, pacing, and editing. Future research should examine emotionality flips in purely audio-based formats (e.g., Spotify podcasts), other narrative forms (e.g., vlogs), and multimodal or short-form content (e.g., TikTok, Instagram Reels). Understanding how emotional dynamics manifest beyond text is increasingly important as platforms evolve.

DECLARATIONS

The authors declare that they have no conflict of interest.

REFERENCES

Adaval, R., & Wyer Jr., R. S. (1998). The Role of Narratives in Consumer Information Processing. *Journal of Consumer Psychology*, *7*(3), 207–245. https://doi.org/10.1207/s15327663jcp0703_01

Adobe Express. (2024, December 7). 9 tested tips to boost your Instagram engagement in 2025. https://www.adobe.com/express/learn/blog/instagram-engagement

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the Tribes of Fluency to Form a Metacognitive Nation. *Personality and Social Psychology Review*, *13*(3), 219–235. https://doi.org/10.1177/1088868309341564

Appel, M., Gnambs, T., & Maio, G. R. (2012). A Short Measure of the Need for Affect. *Journal of Personality Assessment*, *94*(4), 418–426. https://doi.org/10.1080/00223891.2012.666921

Argo, J. J., Zhu, R., & Dahl, D. W. (2008). Fact or Fiction: An Investigation of Empathy Differences in Response to Emotional Melodramatic Entertainment. *Journal of Consumer Research*, *34*(5), 614–623. https://doi.org/10.1086/521907

Backlinko Team. (2024, January 29). 13 Podcast Statistics You Need To Know. https://backlinko.com/podcast-stats#podcast-growth

Berger, J. (2011). Arousal Increases Social Transmission of Information. *Psychological Science*, *22*(7), 891–893. https://doi.org/10.1177/0956797611413294

Berger, J. (2014). Word of mouth and interpersonal communication: A review and directions for future research. *Journal of Consumer Psychology*, *24*(4), 586–607. https://doi.org/10.1016/j.jcps.2014.05.002

Berger, J., Kim, Y. D., & Meyer, R. (2021). What Makes Content Engaging? How Emotional Dynamics Shape Success. *Journal of Consumer Research*, *48*(2), 235–250. https://doi.org/10.1093/jcr/ucab010

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

Berger, J., Moe, W. W., & Schweidel, D. A. (2023). What Holds Attention? Linguistic Drivers of Engagement. *Journal of Marketing*, *87*(5), 793–809. https://doi.org/10.1177/00222429231152880

Berlyne, D. E. (1966). Curiosity and Exploration. *Science*, *153*(3731), 25–33. https://doi.org/10.1126/science.153.3731.25

Bestgen, Y. (1994). Can emotional valence in stories be determined from words? *Cognition and Emotion*, *8*(1), 21–36. https://doi.org/10.1080/02699939408408926

Blascovich, J. (1990). Individual Differences in Physiological Arousal and Perception of Arousal: Missing Links in Jamesian Notions of Arousal-Based Behaviors. *Personality and Social Psychology Bulletin*, *16*(4), 665–675. https://doi.org/10.1177/0146167290164007

Boyd, R. L., Blackburn, K. G., & Pennebaker, J. W. (2020). The narrative arc: Revealing core narrative structures through text analysis. *Science Advances*, *6*(32), eaba2196. https://doi.org/10.1126/sciadv.aba2196

Campbell, M. C., & Kirmani, A. (2000). Consumers' Use of Persuasion Knowledge: The Effects of Accessibility and Cognitive Capacity on Perceptions of an Influence Agent. *Journal of Consumer Research*, *27*(1), 69–83. https://doi.org/10.1086/314309

Cascio Rizzo, G. L., Berger, J., De Angelis, M., & Pozharliev, R. (2023). How Sensory Language Shapes Influencer's Impact. *Journal of Consumer Research*, *50*(4), 810–825. https://doi.org/10.1093/jcr/ucad017

Cascio Rizzo, G. L., Villarroel Ordenes, F., Pozharliev, R., De Angelis, M., & Costabile, M. (2023). How High-Arousal Language Shapes Micro- Versus Macro-Influencers' Impact. *Journal of Marketing*, 00222429231207636. https://doi.org/10.1177/00222429231207636

Chatman, S. (1978). *Story and Discourse: Narrative Structure in Fiction and Film*. Ithaca: NY: Cornell University Press.

Chen, Z., & Berger, J. (2013). When, Why, and How Controversy Causes Conversation. *Journal of Consumer Research*, *40*(3), 580–593. https://doi.org/10.1086/671465

Chung, J. (Jae), Ding, Y., & Kalra, A. (2023). I Really Know You: How Influencers Can Increase Audience Engagement by Referencing Their Close Social Ties. *Journal of Consumer Research*, *50*(4), 683–703. https://doi.org/10.1093/jcr/ucad019

Consiglio, I., De Angelis, M., & Costabile, M. (2018). The Effect of Social Density on Word of Mouth. *Journal of Consumer Research*. https://doi.org/10.1093/jcr/ucy009

Cox, M., Harrison, H. L., Partelow, S., Curtis, S., Elser, S. R., Hammond Wagner, C., et al. (2023). How academic podcasting can change academia and its relationship with society: A conversation and guide. *Frontiers in Communication*, *8*. https://www.frontiersin.org/articles/10.3389/fcomm.2023.1090112

Cramer-Flood, E. (2021, October 22). Global Podcast Listener Forecast 2021–2025. https://www.emarketer.com/content/global-podcast-listener-forecast-2021-2025

Culler, J. D. (1981). *The Pursuit of Signs: Semiotics, Literature, Deconstruction* (Ithaca.). NY: Cornell University Press.

Devereux, E., Grimmer, L., & Grimmer, M. (2020). Consumer engagement on social media: Evidence from small retailers. *Journal of Consumer Behaviour*, *19*(2), 151–159. https://doi.org/10.1002/cb.1800

Duarte, F. (2023a, December 5). TikTok Content Creator Statistics (2024).

Duarte, F. (2023b, December 5). YouTube Content Creator Statistics (2024). https://explodingtopics.com/blog/youtube-creator-stats

Dubois, D., Bonezzi, A., & De Angelis, M. (2016). Sharing with Friends versus Strangers: How Interpersonal Closeness Influences Word-of-Mouth Valence. *Journal of Marketing Research*, *53*(5), 712–727. https://doi.org/10.1509/jmr.13.0312

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). Emotion in the human face: Guidelines for research and an integration of findings. *Emotion in the human face: Guidelines for research and an integration of findings.*, xii, 191–xii, 191.

Escalas, J. E. (2004a). Imagine Yourself in the Product: Mental Simulation, Narrative Transportation, and Persuasion. *Journal of Advertising*, *33*(2), 37–48. https://doi.org/10.1080/00913367.2004.10639163

Escalas, J. E. (2004b). Narrative processing: Building consumer connections to brands. *Journal of Consumer Psychology*, *14*(1–2), 168–180. https://doi.org/10.1207/s15327663jcp1401&2_19

Escalas, J. E. (2007). Self-Referencing and Persuasion: Narrative Transportation versus Analytical Elaboration. *Journal of Consumer Research*, *33*(4), 421–429. https://doi.org/10.1086/510216

Feiereisen, S., Rasolofoarison, D., Russell, C. A., & Schau, H. J. (2021). One Brand, Many Trajectories: Narrative Navigation in Transmedia. *Journal of Consumer Research*, *48*(4), 651–681. https://doi.org/10.1093/jcr/ucaa046

Fischer, S. (2025, April 29). Axios Media Trends. https://www.axios.com/newsletters/axios-media-trends-da7be500-2458-11f0-b89e-d1e9e0765a00

Fisk, J. E. (2002). Judgments under uncertainty: Representativeness or potential surprise? *British Journal of Psychology*, *93*(4), 431–449. https://doi.org/10.1348/000712602761381330

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*(3), 221–233. https://doi.org/10.1037/h0057532

Forster, M., Leder, H., & Ansorge, U. (2016). Exploring the Subjective Feeling of Fluency. *Experimental Psychology*, *63*(1), 45–58. https://doi.org/10.1027/1618-3169/a000311

Garcia, C. (2021, September 10). How to engage your audience. https://creators.google/en-us/content-creation-guides/audience-engagement/everything-you-need-to-know-about-audience-engagement/

Gardner, J., & Lehnert, K. (2016). What’s new about new media? How multi-channel networks work with content creators. *Business Horizons*, *59*(3), 293–302. https://doi.org/10.1016/j.bushor.2016.01.009

Garfinkel, S. N., Minati, L., Gray, M. A., Seth, A. K., Dolan, R. J., & Critchley, H. D. (2014). Fear from the Heart: Sensitivity to Fear Stimuli Depends on Individual Heartbeats. *The Journal of Neuroscience*, *34*(19), 6573. https://doi.org/10.1523/JNEUROSCI.3507-13.2014

Gee, J. P., & Grosjean, F. (1984). Empirical evidence for narrative structure. *Cognitive Science*, *8*(1), 59–85. https://doi.org/10.1016/S0364-0213(84)80025-7
Goldenberg, J., Lanz, A., Shapira, D., & Stahl, F. (2023). Targeting Nearby Influencers: The Acceleration of Natural Triadic Closure by Leveraging Interconnectors. *Journal of Marketing*, 00222429231223420. https://doi.org/10.1177/00222429231223420
Goldman Sachs. (2023, April 19). The creator economy could approach half-a-trillion dollars by 2027. https://www.goldmansachs.com/intelligence/pages/the-creator-economy-could-approach-half-a-trillion-dollars-by-2027.html
Green, M., & Brock, T. (2000). The Role of Transportation in the Persuasiveness of Public Narrative. *Journal of personality and social psychology*, *79*, 701–21. https://doi.org/10.1037/0022-3514.79.5.701
Green, M. C. (2008). Transportation theory. *The international encyclopedia of communication*.
Gu, X., Zhang, X., & Kannan, P. K. (2024). Influencer Mix Strategies in Livestream Commerce: Impact on Product Sales. *Journal of Marketing*, *88*(4), 64–83. https://doi.org/10.1177/00222429231213581
Haddock, G., & Maio, G. R. (2019). Inter-individual differences in attitude content: Cognition, affect, and attitudes. *Advances in experimental social psychology, Vol. 59*, 53–102. https://doi.org/10.1016/bs.aesp.2018.10.002
Hamby, A., & Brinberg, D. (2016). Happily Ever After: How Ending Valence Influences Narrative Persuasion in Cautionary Stories. *Journal of Advertising*, *45*(4), 498–508. https://doi.org/10.1080/00913367.2016.1262302
Hamby, A., & Escalas, J. E. (2024). Connecting the Plot Points: How Consumers Use and Respond to Narratives. *Journal of the Association for Consumer Research*, *9*(1), 1–8. https://doi.org/10.1086/727829
Hartmann, J., Heitmann, M., Schamp, C., & Netzer, O. (2021). The Power of Brand Selfies. *Journal of Marketing Research*, *58*(6), 1159–1177. https://doi.org/10.1177/00222437211037258
Hayes, A. F. (2022). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (third edition.). New York, NY: The Guilford Press.
Heilman, K. M. (1997). The neurobiology of emotional experience. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *9*(3), 439–448. https://doi.org/10.1176/jnp.9.3.439
Herzenstein, M., Sonenshein, S., & Dholakia, U. M. (2011). Tell Me a Good Story and I May Lend you Money: The Role of Narratives in Peer-to-Peer Lending Decisions. *Journal of Marketing Research*, *48*(SPL), S138–S149. https://doi.org/10.1509/jmkr.48.SPL.S138
Hough, Q. (2019, March 15). Why The Red Wedding Is Still Game of Thrones' Most Shocking Moment. https://screenrant.com/red-wedding-game-thrones-shocking-moment/
Howarth, J. (2024, May 9). 30+ Incredible Creator Economy Statistics (2024). https://explodingtopics.com/blog/creator-economy-stats
Huang, M.-H., & Rust, R. T. (2025). The GenAI Future of Consumer Research. *Journal of Consumer Research*, *52*(1), 4–17. https://doi.org/10.1093/jcr/ucaf013
Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media* (Vol. 8, pp. 216–225).
Jiang, H., Messinger, P. R., Liu, Y., Lu, Z., Yang, S., & Li, G. (2024). Divergent Versus Relevant Ads: How Creative Ads Affect Purchase Intention for New Products.

Journal of Marketing Research, *61*(2), 271–289. https://doi.org/10.1177/00222437231187630

Kim, E., Shoenberger, H., Kim, D., Thorson, E., & Zihang, E. (2024). Novelty vs. trust in virtual influencers: exploring the effectiveness of human-like virtual influencers and anime-like virtual influencers. *International Journal of Advertising*, 1–31. https://doi.org/10.1080/02650487.2024.2386916

KLERES, J. (2011). Emotions and Narrative Analysis: A Methodological Approach. *Journal for the Theory of Social Behaviour*, *41*(2), 182–202. https://doi.org/10.1111/j.1468-5914.2010.00451.x

Kovacheva, A., Nikolova, H., & Lamberton, C. (2022). Will he buy a surprise? Gender differences in the purchase of surprise offerings. *Journal of Retailing*, *98*(4), 667–684. https://doi.org/10.1016/j.jretai.2022.04.002

Kozinets, R. V., Gretzel, U., & Gambetti, R. (2023). *Influencers and Creators: Business, Culture and Practice*. Sage Publications.

Krause-Galoni, R., & Rucker, D. D. (2024). Fact versus Fiction: Blending Storytelling with Product Attributes in Advertisements. *Journal of the Association for Consumer Research*, *9*(1), 58–70. https://doi.org/10.1086/727835

Kumar, N. (2025, June 25). 32+ Creator Economy Statistics of 2025 (Market Size Data). https://www.demandsage.com/creator-economy-statistics/

Lacka, E., Boyd, D. E., Ibikunle, G., & Kannan, P. K. (2022). Measuring the Real-Time Stock Market Impact of Firm-Generated Content. *Journal of Marketing*, *86*(5), 58–78. https://doi.org/10.1177/00222429211042848

Lanz, A., Goldenberg, J., Shapira, D., & Stahl, F. (2024). Buying Future Endorsements from Prospective Influencers on User-Generated Content Platforms. *Journal of Marketing Research*, 00222437231207323. https://doi.org/10.1177/00222437231207323

Lee, J. K. (2021). Emotional Expressions and Brand Status. *Journal of Marketing Research*, *58*(6), 1178–1196. https://doi.org/10.1177/00222437211037340

Lei, A., Willems, R. M., & Eekhof, L. S. (2023). Emotions, fast and slow: processing of emotion words is affected by individual differences in need for affect and narrative absorption. *Cognition and Emotion*, *37*(5), 997–1005. https://doi.org/10.1080/02699931.2023.2216445

Leung, F. F., Gu, F. F., Li, Y., Zhang, J. Z., & Palmatier, R. W. (2022). Influencer Marketing Effectiveness. *Journal of Marketing*, *86*(6), 93–115. https://doi.org/10.1177/00222429221102889

Leung, F. F., Gu, F. F., & Palmatier, R. W. (2022). Online influencer marketing. *Journal of the Academy of Marketing Science*, *50*(2), 226–251. https://doi.org/10.1007/s11747-021-00829-4

Li, C., Xie, M., & Wang, M. (2021). Understanding User Engagement in Online Communities during COVID-19 Pandemic: Evidence from Sentiment and Semantic Analysis on YouTube.

Lin, Y., Yao, D., & Chen, X. (2021). Happiness Begets Money: Emotion and Engagement in Live Streaming. *Journal of Marketing Research*, *58*(3), 417–438. https://doi.org/10.1177/00222437211002477

Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual differences in the motivation to approach or avoid emotions. *Journal of Personality*, *69*(4), 583–615. https://doi.org/10.1111/1467-6494.694156

Mandler, G. (1982). The Structure of Value: Accounting for Taste. In M. S. Clark & S. T. Fiske (Eds.), *Affect and Cognition: The Seventeenth Annual Carnegie Symposium on Cognition* (pp. 3–36). Hillsdale, NJ: Erlbaum.

Mardon, R., Cocker, H., & Daunt, K. (2023). How Social Media Influencers Impact Consumer Collectives: An Embeddedness Perspective. *Journal of Consumer Research*, *50*(3), 617–644. https://doi.org/10.1093/jcr/ucad003

Markowitz, D. M. (2023). Instrumental Goal Activation Increases Online Petition Support Across Languages. *Journal of personality and social psychology*, *124*(6), 1133–1145. https://doi.org/10.1037/pspa0000333

Markowitz, D. M., & Shulman, H. C. (2021). The predictive utility of word familiarity for online engagements and funding. *Proceedings of the National Academy of Sciences*, *118*(18), e2026045118. https://doi.org/10.1073/pnas.2026045118

Marsh, E., & Fazio, L. (2006). Learning errors from fiction: Difficulties in reducing reliance on fictional stories. *Memory & cognition*, *34*, 1140–9. https://doi.org/10.3758/BF03193260

Melumad, S., Inman, J. J., & Pham, M. T. (2019). Selectively Emotional: How Smartphone Use Changes User-Generated Content. *Journal of Marketing Research*, *56*(2), 259–275. https://doi.org/10.1177/0022243718815429

Milkman, K. L., & Berger, J. (2014). The science of sharing and the sharing of science. *Proceedings of the National Academy of Sciences*, *111*(supplement_4), 13642–13649. https://doi.org/10.1073/pnas.1317511111

Morin, O., & Acerbi, A. (2017). Birth of the cool: a two-centuries decline in emotional expression in Anglophone fiction. *Cognition and Emotion*, *31*(8), 1663–1675. https://doi.org/10.1080/02699931.2016.1260528

Mosley, B., Schweidel, D. A., & Zhang, K. (2024). When Connection Turns to Anger: How Consumer–Brand Relationship and Crisis Type Moderate Language on Social Media. *Journal of Consumer Research*, *50*(5), 907–922. https://doi.org/10.1093/jcr/ucad027

NELL, V. (1988). *Lost in a Book*. Yale University Press. https://doi.org/10.2307/j.ctt1ww3vk3

Nielsen, J. H., & Escalas, J. E. (2010). Easier is not always better: The moderating role of processing type on preference fluency. *Journal of Consumer Psychology*, *20*(3), 295–305. https://doi.org/10.1016/j.jcps.2010.06.016

Noh, S. R., & Isaacowitz, D. M. (2013). Emotional faces in context: Age differences in recognition accuracy and scanning patterns. *Emotion (Washington, D.C.)*, *13*(2), 238–249. https://doi.org/10.1037/a0030234

Noseworthy, T. J., Di Muro, F., & Murray, K. B. (2014). The Role of Arousal in Congruity-Based Product Evaluation. *Journal of Consumer Research*, *41*(4), 1108–1126. https://doi.org/10.1086/678301

Packard, G., Berger, J., & Boghrati, R. (2023). How Verb Tense Shapes Persuasion. *Journal of Consumer Research*, *50*(3), 645–660. https://doi.org/10.1093/jcr/ucad006

Pancer, E., Chandler, V., Poole, M., & Noseworthy, T. J. (2019). How Readability Shapes Social Media Engagement. *Journal of Consumer Psychology*, *29*(2), 262–270. https://doi.org/10.1002/jcpy.1073

Papies, D., Ebbes, P., & van Heerde, H. (2017). Addressing Endogeneity in Marketing Models (pp. 581–627). https://doi.org/10.1007/978-3-319-53469-5_18

Park, S., & Gupta, S. (2012). Handling Endogenous Regressors by Joint Estimation Using Copulas. *Marketing Science*, *31*(4), 567–586.

Peres, R., Schreier, M., Schweidel, D. A., & Sorescu, A. (2024). The creator economy: An introduction and a call for scholarly research. *International Journal of Research in Marketing*, *41*(3), 403–410. https://doi.org/10.1016/j.ijresmar.2024.07.005

Perse, E. (1998). Implications of cognitive and affective involvement for channel changing. *Journal of Communication*, *48*(3), 49–68. https://doi.org/10.1111/j.1460-2466.1998.tb02759.x

Peterson, T. (2024, July 15). Why longer videos are becoming more commonplace on YouTube. https://digiday.com/future-of-tv/why-longer-videos-are-becoming-more-commonplace-on-youtube/

Pham, M. T. (1998). Representativeness, Relevance, and the Use of Feelings in Decision Making. *Journal of Consumer Research*, *25*(2), 144–159. https://doi.org/10.1086/209532

Ptaszynski, M., Dokoshi, H., Oyama, S., Rzepka, R., Kurihara, M., Araki, K., & Momouchi, Y. (2013). Affect analysis in context of characters in narratives. *Expert Systems with Applications*, *40*(1), 168–176. https://doi.org/10.1016/j.eswa.2012.07.025

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*(5), 379–399. https://doi.org/10.1037/h0046234

Schuster, S., Hawelka, S., Hutzler, F., Kronbichler, M., & Richlan, F. (2016). Words in Context: The Effects of Length, Frequency, and Predictability on Brain Responses During Natural Reading. *Cerebral Cortex*, *26*(10), 3889–3904. https://doi.org/10.1093/cercor/bhw184

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition and Emotion*, *14*(4), 433–440. https://doi.org/10.1080/026999300402745

Seiler, S., Yao, S., & Wang, W. (2017). Does Online Word of Mouth Increase Demand? (And How?) Evidence from a Natural Experiment. *Marketing Science*, *36*(6), 838–861. https://doi.org/10.1287/mksc.2017.1045

Sepehri, A., Duclos, R., & Haghighi, N. (2025). Ideas worth spreading? When, how, and for whom information load hurts online talks' popularity. *Journal of Personality and Social Psychology*, *128*(2), 281–299. https://doi.org/10.1037/pspa0000430

Sepehri, A., Markowitz, D. M., & Duclos, R. (2021). The Location of Maximum Emotion in Deceptive and Truthful Texts. *Social Psychological and Personality Science*, *12*(6), 996–1004. https://doi.org/10.1177/1948550620949730

Shuster, M. M., Mikels, J. A., & Camras, L. A. (2017). Adult age differences in the interpretation of surprised facial expressions. *Emotion*, *17*(2), 191–195. https://doi.org/10.1037/emo0000234

Somervail, R., Zhang, F., Novembre, G., Bufacchi, R. J., Guo, Y., Crepaldi, M., et al. (2021). Waves of Change: Brain Sensitivity to Differential, not Absolute, Stimulus Intensity is Conserved Across Humans and Rats. *Cerebral Cortex*, *31*(2), 949–960. https://doi.org/10.1093/cercor/bhaa267

Spiller, S. A., Fitzsimons, G. J., Lynch, J. G., & Mcclelland, G. H. (2013). Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression. *Journal of Marketing Research*, *50*(2), 277–288. https://doi.org/10.1509/jmr.12.0420

Spotify. (2024, July 29). How to grow your podcast audience. https://creators.spotify.com/resources/grow/grow-audience

Stroud, N. J., & Muddiman, A. (2019). Social Media Engagement With Strategy- and Issue-Framed Political News. *Journal of Communication*, *69*(5), 443–466. https://doi.org/10.1093/joc/jqz029

Teigen, K. H., & Keren, G. (2003). Surprises: low probabilities or high contrasts? *Cognition*, *87*(2), 55–71. https://doi.org/10.1016/s0010-0277(02)00201-9

Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-Induced Engagement in Internet Video Advertisements. *Journal of Marketing Research*, *49*(2), 144–159. https://doi.org/10.1509/jmr.10.0207

Tellis, G. J., MacInnis, D. J., Tirunillai, S., & Zhang, Y. (2019). What Drives Virality (Sharing) of Online Digital Content? The Critical Role of Information, Emotion, and

Brand Prominence. *Journal of Marketing*, *83*(4), 1–20. https://doi.org/10.1177/0022242919841034
Torta, D., Ninghetto, M., Ricci, R., & Legrain, V. (2020). Rating the Intensity of a Laser Stimulus, but Not Attending to Changes in Its Location or Intensity Modulates the Laser-Evoked Cortical Activity. *Frontiers in Human Neuroscience*, *14*. https://doi.org/10.3389/fnhum.2020.00120
Toubia, O., Berger, J., & Eliashberg, J. (2021). How quantifying the shape of stories predicts their success. *Proceedings of the National Academy of Sciences*, *118*(26), e2011695118. https://doi.org/10.1073/pnas.2011695118
Toubia, O., Gui, G. Z., Peng, T., Merlau, D. J., Li, A., & Chen, H. (2025). Database Report: Twin-2K-500: A Data Set for Building Digital Twins of over 2,000 People Based on Their Answers to over 500 Questions. *Marketing Science*. https://doi.org/10.1287/mksc.2025.0262
Valenzuela, A., Mellers, B., & Strebel, J. (2010). Pleasurable Surprises: A Cross-Cultural Study of Consumer Responses to Unexpected Incentives. *Journal of Consumer Research*, *36*, 792–805. https://doi.org/10.1086/605592
Van Boven, L., Kane, J., Mcgraw, A. P., & Dale, J. (2010). Feeling Close: Emotional Intensity Reduces Perceived Psychological Distance. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1531661
van Laer, T., de Ruyter, K., Visconti, L. M., & Wetzels, M. (2014). The Extended Transportation-Imagery Model: A Meta-Analysis of the Antecedents and Consequences of Consumers' Narrative Transportation. *Journal of Consumer Research*, *40*(5), 797–817. https://doi.org/10.1086/673383
van Laer, T., Edson Escalas, J., Ludwig, S., & van den Hende, E. A. (2019). What Happens in Vegas Stays on TripAdvisor? A Theory and Technique to Understand Narrativity in Consumer Reviews. *Journal of Consumer Research*, *46*(2), 267–285. https://doi.org/10.1093/jcr/ucy067
Whiteford, K. L., & Oxenham, A. J. (2023). Sensitivity to Frequency Modulation is Limited Centrally. *The Journal of Neuroscience*, *43*(20), 3687. https://doi.org/10.1523/JNEUROSCI.0995-22.2023
Yin, D., Bond, S. D., & Zhang, H. (2017). Keep Your Cool or Let it Out: Nonlinear Effects of Expressed Arousal on Perceptions of Consumer Reviews. *Journal of Marketing Research*, *54*(3), 447–463. https://doi.org/10.1509/jmr.13.0379
Yuan, H., Che, Z., Li, S., Zhang, Y., Hu, X., & Luo, S. (2024). The high dimensional psychological profile and cultural bias of ChatGPT. *ArXiv*, *abs/2405.03387*. https://api.semanticscholar.org/CorpusID:269605350

How Should Content Creators Narrate Their Content? The Impact of Emotionality Flips on Audience Engagement

WEB APPENDIX

Contents

Web Appendix A: Literature Review on Emotionality in Content and Narrativity

Table A1. Summary of representative research on emotionality in content

Note: API = application programming interface. eWOM = electronic word of mouth. LIWC = Linguistic Inquiry and Word Count. SD = standard deviation. UGC = User-generated content. NRC = a list of English words and their associations with eight basic emotions. PCA = Principal components analysis.

Authors	Context and Method	Emotionality Measurement	Main Findings
Mosley et al. (2024)	*Context*: Consumer complaint commented on Facebook *Method*: Event study using LIWC and NRC	Linguistic attributes of anger	Following values-related brand failures, consumers who have previously interacted with the brand express more anger than those who have not.
Berger et al. (2023)	*Context:* Reading sessions *Method:* Natural language processing (LIWC and Mohammad and Bravo-Marquez (2017) method) and	Emotional valence and language linked to specific emotions	Processing ease and emotion play a crucial role in capturing attention.
Ravichandran & Deng (2023)	*Context:* Managerial responses to consumer complaints on Tripadvisor.com *Method:* Natural language processing and deep learning	Positive vs. negative reviews	Responding to customer complaints about interactional unfairness with emotional cues increases future review valence.
Berger et al. (2021)	*Context:* Movies and online articles *Method*: Automated sentiment analysis, controlled	Measuring SDs in the sentiment of adjoining chunks	Period-to-period shifts in sentiment valence enhance engagement.
Lee, (2021)	*Context*: Text and images posted by brands on Twitter and Instagram *Method:* Text analysis, computer vision, controlled experiments (LIWC)	Average emotionality	There is a negative correlation between brand status and the emotionality expressed in the brand's social media posts.
Lin et al. (2021)	*Context:* Live streams *Method:* Deep neural network and text analysis (LIWC)	Broadcasters' and audiences' emotions at the minute level	Happier broadcasters make the audience happier, and smiles can increase engagement.
Li & Xie (2020)	*Context:* Twitter and Instagram data sets *Method*: Google Cloud Vision API and text analysis (LIWC)	Face detection and emotional state	The mere presence of image content increases user engagement.

Berman et al., (2019)	*Context:* Tweets *Method*: Automated natural language processing (LIWC, Speciteller, Visual content)	1) The use of positive versus negative emotional words 2) The use of words referencing achievement, power, and reward 3) Composite indices of analytical and authentic writing	Emotionality in a tweet can enhance retweets during and after a political debate.
van Laer et al. (2019)	*Context:* TripAdvisor reviews *Method*: Text analysis (LIWC), controlled experiments	Emotionality intensity at the sentence level	Reviews with more emotionally varied genres are more persuasive.
Tellis et al. (2019)	*Context:* YouTube video ads *Method*: PCA with Varimax rotation, human coders	Labeled positive components of emotions and used in the empirical analysis	Emotional ads are shared more on general platforms (Facebook, Google+, Twitter) than on LinkedIn. Conversely, informational ads are more commonly shared on LinkedIn.
Herhausen et al. (2019)	*Context*: Negative customer posts on Facebook, and firm responses to them *Method*: Text analysis (LIWC)	Intensity of positive- and negative-emotion words in each post	Negative eWOM messages with higher emotionality are associated with greater virality.
Melumad et al. (2019)	*Context*: UGC with mobile devices vs. personal computers *Method*: Text analysis (LIWC), human coders	Emotionality and emotional valence	People tend to generate brief content on their phones; therefore, they use less emotional language.
Ransbotham et al. (2019)	*Context*: Reviews written with mobile vs. nonmobile devices on http://urbanspoon.com/ *Method*: Text analysis (LIWC)	Average emotionality	Users with mobile devices express more emotional WOM.
Akpinar & Berger (2017)	*Context*: Online ads *Method*: Human coders and laboratory experiment	Coding: Informative vs. Emotional	Emotional appeals (which use drama, mood, music, and other emotion-eliciting strategies) are more likely to be shared than informative appeals.
Yazdani et al. (2018)	*Context*: Amazon reviews and sales activity for new music albums *Method*: Text analysis (LIWC)	The extent of both positive and negative emotional content in the	Emotionality in reviews can enhance Amazon sales.

Heimbach & Hinz (2016)	*Context:* Sample of German articles *Method*: Automated sentiment analysis (LIWC)	Percentage of all positive and negative words in the article	The relationship between emotional positivity and virality is an inverted U shape.
Berger & Milkman (2012)	*Context*: New York Times articles published over a three-month period *Method*: Automated sentiment analysis (LIWC), human coders,	Percentage of all positive and negative words in the article	High-arousal emotions like awe, anger, or anxiety make content more viral, while low-arousal emotions like sadness make it less viral.

Table A2. Summary of representative research on narrativity in content

Authors	Context	Contribution	Main Findings
van Laer et al. (2019)	Online reviews	Narrative content, discourse, and persuasion	Reviews with more emotionally varied genres are more persuasive.
Phillips and McQuarrie (2010)	Fashion ads	Narrative transportation and persuasion	Grotesque imagery in fashion ads can elicit narrative transportation and serve as an indirect route to persuasion.
Adaval and Wyer Jr. (1998)	Vacation brochures	Narrative discourse	Narrative presentation of product information (vs. list format) leads to more favorable judgments.
Adaval, Isbell, and Wyer (2007)	Life events	Narrative discourse	Pictures enhance impression formation when events are conveyed as a coherent narrative, but hinder it when presented in a list format.
Martin et al. (2024)	Brand history	Narrative content and persuasion	Consumers form brand heritage perceptions through narrative processing.
Valenzuela and Galli (2024)	Online reviews	Narrative content, transportation, and persuasion	Personal narratives increase transportation and enhance persuasiveness by eliciting a sense of connection and mental simulation.
Bublitz et al. (2024)	Individual and collective activist stories	Narrative content	Proposes a conceptual framework distinguishing between individual and collective narratives to explain how storytelling fosters unity and action in social movements.
Green (2004)	Fictitious story	Narrative discourse, transportation, and persuasion	Transportation is enhanced by prior knowledge and perceived realism, leading to a greater endorsement of story-consistent beliefs.

Escalas (2007)	ads	Narrative contents, transportation, and persuasion	Narrative (vs. analytical) self-referencing leads to persuasion via transportation and is robust to weak arguments.
van Laer et al. (2014)	Meta-analyses	Narrative content, discourse, transportation, and persuasion	Narrative transportation is a robust mechanism driving persuasion and memory.
Green and Brock (2000)	Fictitious and factual stories	Narrative discourse, transportation, and persuasion	Transporting a narrative into a story enhances story-consistent beliefs and favorable evaluations.
Escalas (2004)	ads	Narrative discourse, transportation, and persuasion	Narrative processing may create a link between a brand and the self when consumers attempt to map incoming narrative information into stories in memory.
Krause-Galoni and Rucker (2024)	ads	Narrative transportation and persuasion	Consolidating arguments within a narrative can enhance persuasion.

Web Appendix B: Sample Details

YouTube provides a daily list of popular podcasts at https://www.youtube.com/podcasts/popularshows.

Table B1. Podcast channels and the number of episodes for each channel

Row	Channel	Number of episodes
1	TheMajorityRep	1485
2	ThePatMcAfeeS	1422
3	revolt	1371
4	briantylercohen	1366
5	LawAndCrime	1366
6	LinusTechTips	1298
7	vladtv	1228
8	NoJumper	1224
9	OfficialJimCorne	1176
10	MrRedderYT	1163
11	NBCNews	1116
12	joebuddentv	992
13	60minutes	974
14	InternetTodayTV	940
15	JockoPodcastOff	882
16	PhilipDeFranco	872
17	smoshpit	841
18	LetsReadPodcast	804
19	ClubShayShay	800
20	H3Podcast	736
21	CamNewton	670
22	FreshFitMiami	643
23	lexfridman	632
24	PBDPodcast	629
25	TomBilyeu	615
26	48hours	566
27	BeBustaYT	494
28	annieelise	490
29	TigerBelly	483
30	DrChatterjeeRan	445
31	GeekandSundry	441
32	OfficialFlagrant	433
33	NightcapShow	421
34	BaileySarian	409
35	unsubscribepodc	406
36	MeidasTouch	397
37	comeandtalk2m	378
38	Impaulsive	373
39	JulianDorey	368
40	TheoVon	362
41	TheDiaryOfACEO	286
42	GeorgeJanko	277

43	TheCasualCrimin	272
44	BadFriends	212
45	hubermanlab	205
46	rottenmangopo	184
47	justtrishpodcast	154
48	cancelledwithta	125
49	whatever	113
50	Shane2	59

Table B2. Distributional diagnostics for audience engagement (Studies 1–2)

Study	Skewness	Kurtosis
Study 1	10.65	300.99
Study 2	8.63	108.69
Note. These values exceed the cut points of |1| and |3| for highly skewed (Bulmer 1979) and a highly leptokurtic distribution, respectively (DeCarlo 1997).		

Table B3. Normality tests for emotionality flips (Studies 1 and 2)

Study	Test	Statistic	p-value	Reject normality
Study 1	Shapiro–Wilk	W = 0.86	< .001	✓
	Anderson–Darling	A = 182.66	< .001	✓
	Kolmogorov–Smirnov	D = 0.15	< .001	✓
	Cramér–von Mises	W = 33.56	< .001	✓
Study 2	Shapiro–Wilk	W = 0.96	< .001	✓
	Anderson–Darling	A = ∞	< .001	✓
	Kolmogorov–Smirnov	D = 0.03	< .001	✓
	Cramér–von Mises	W = 147.02	< .001	✓

Figure B1. Study 1: Histogram of audience engagement before (top) and after (bottom) log-transformation

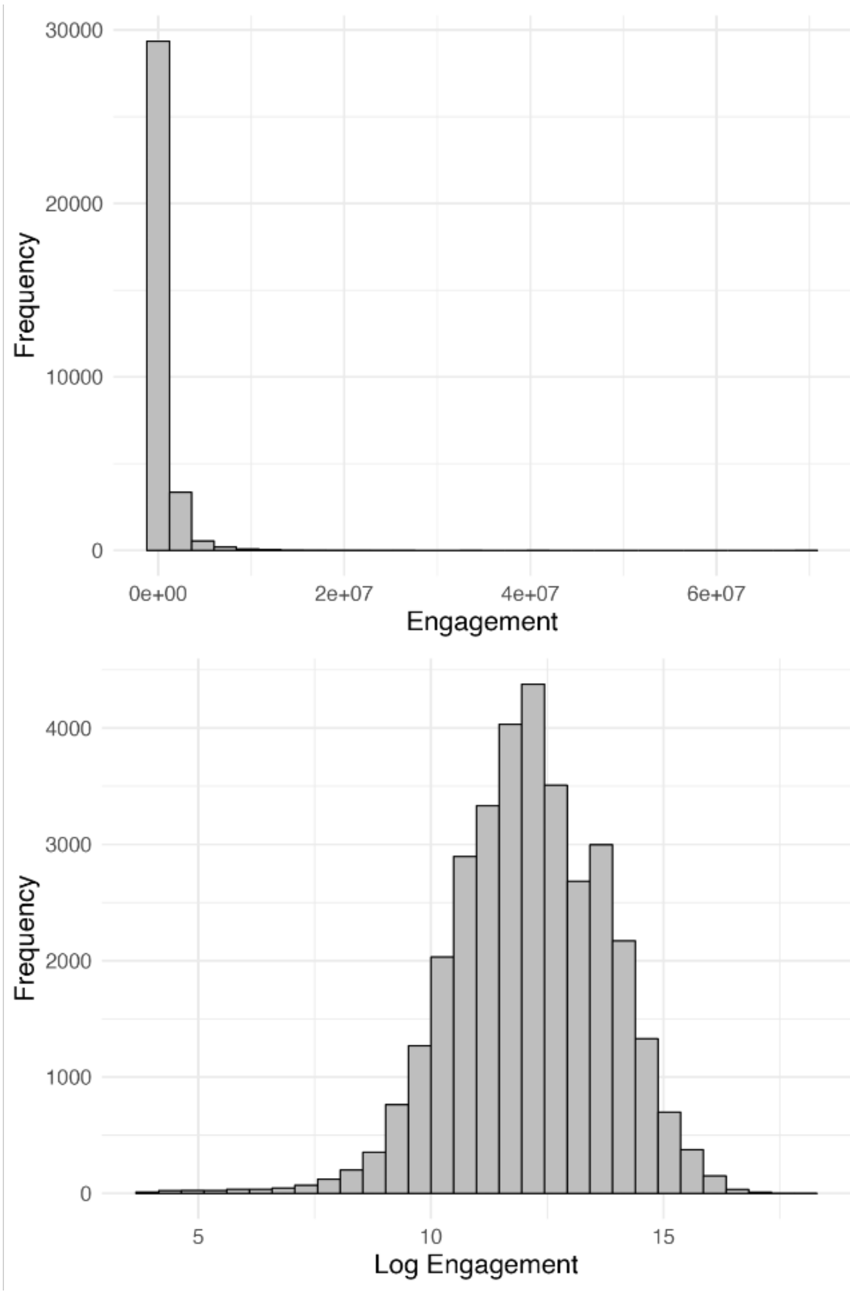


Figure B2. Study 1: Q-Q plot of residuals to assess normality in the regression model

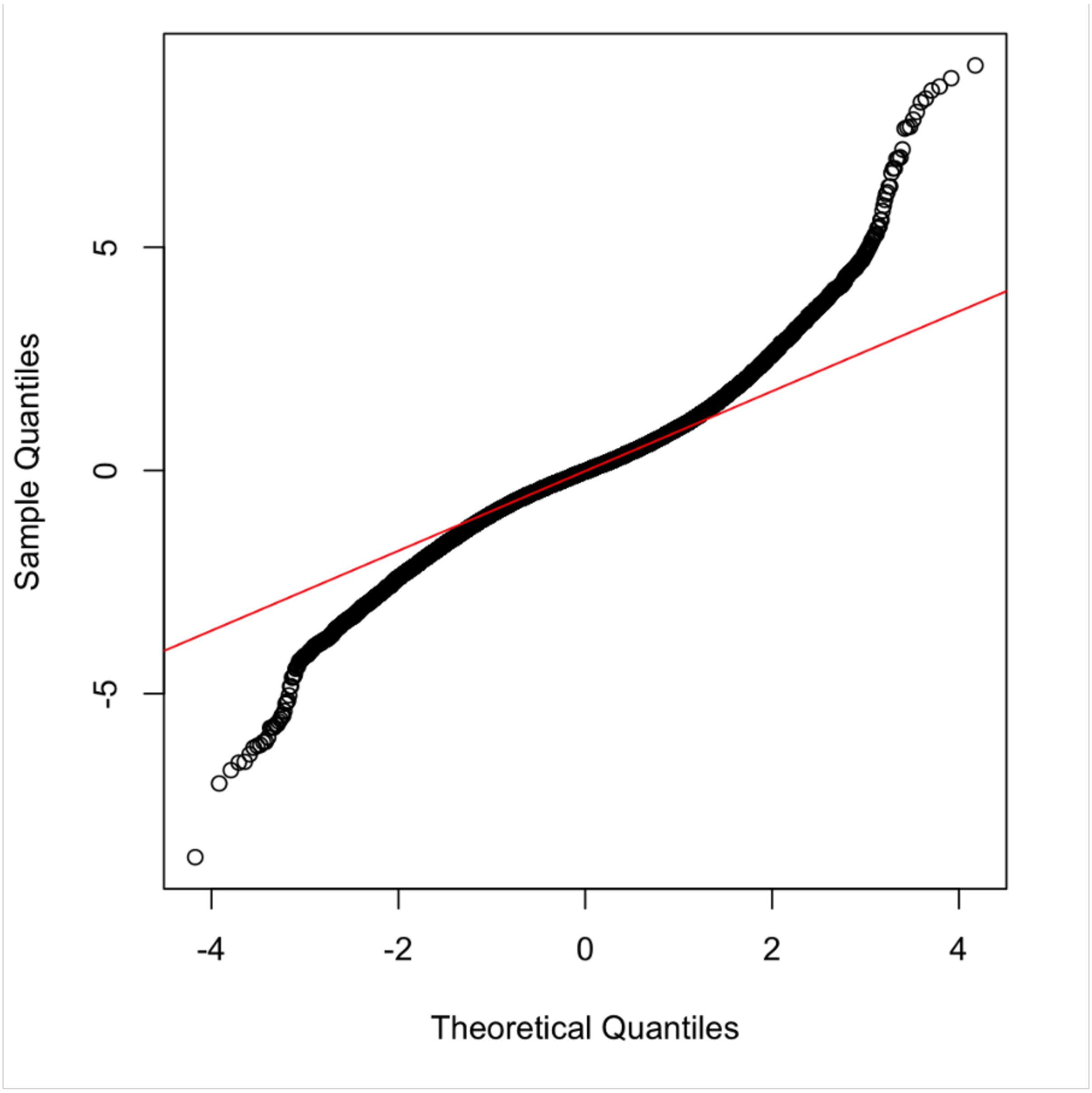


Figure B3. Study 2: Histogram of audience engagement before (top) and after (bottom) log-transformation

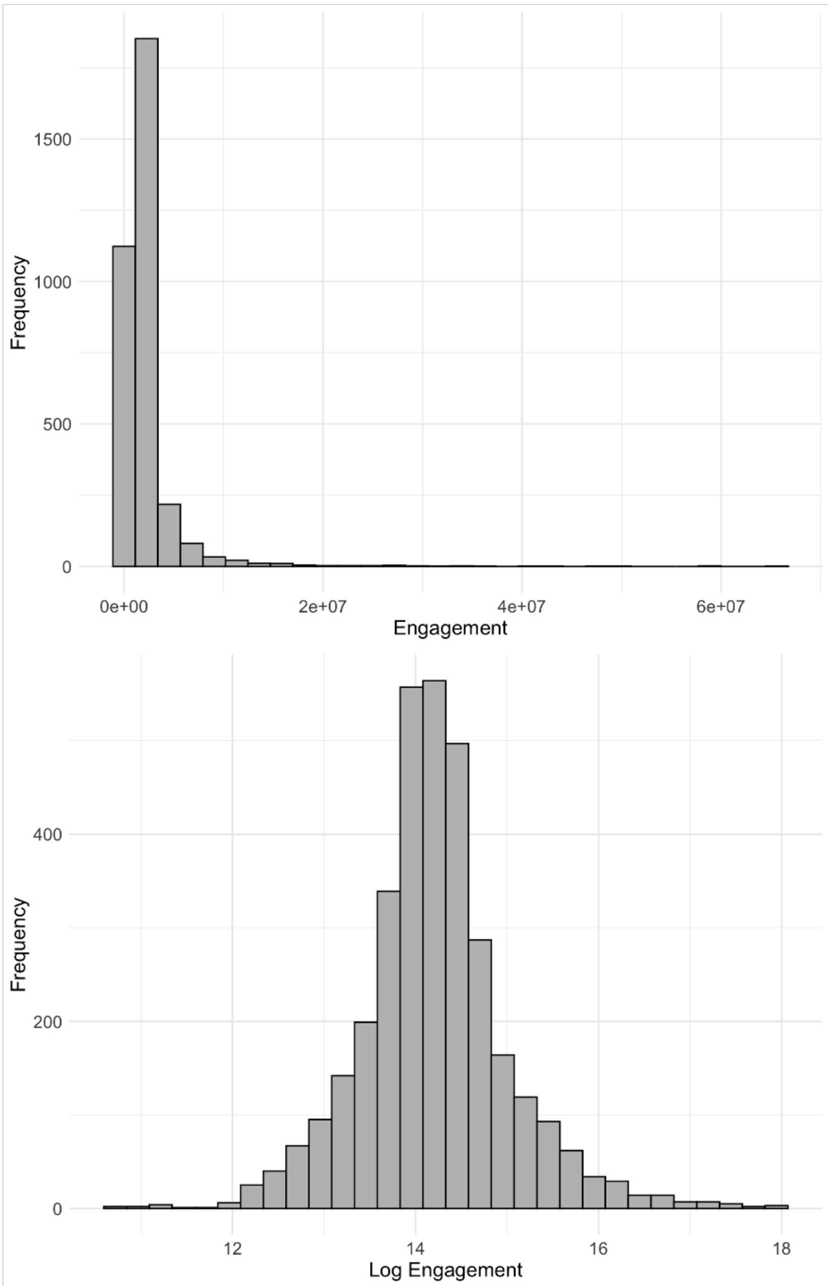


Figure B4. Study 2: Q-Q plot of residuals to assess normality in the regression model

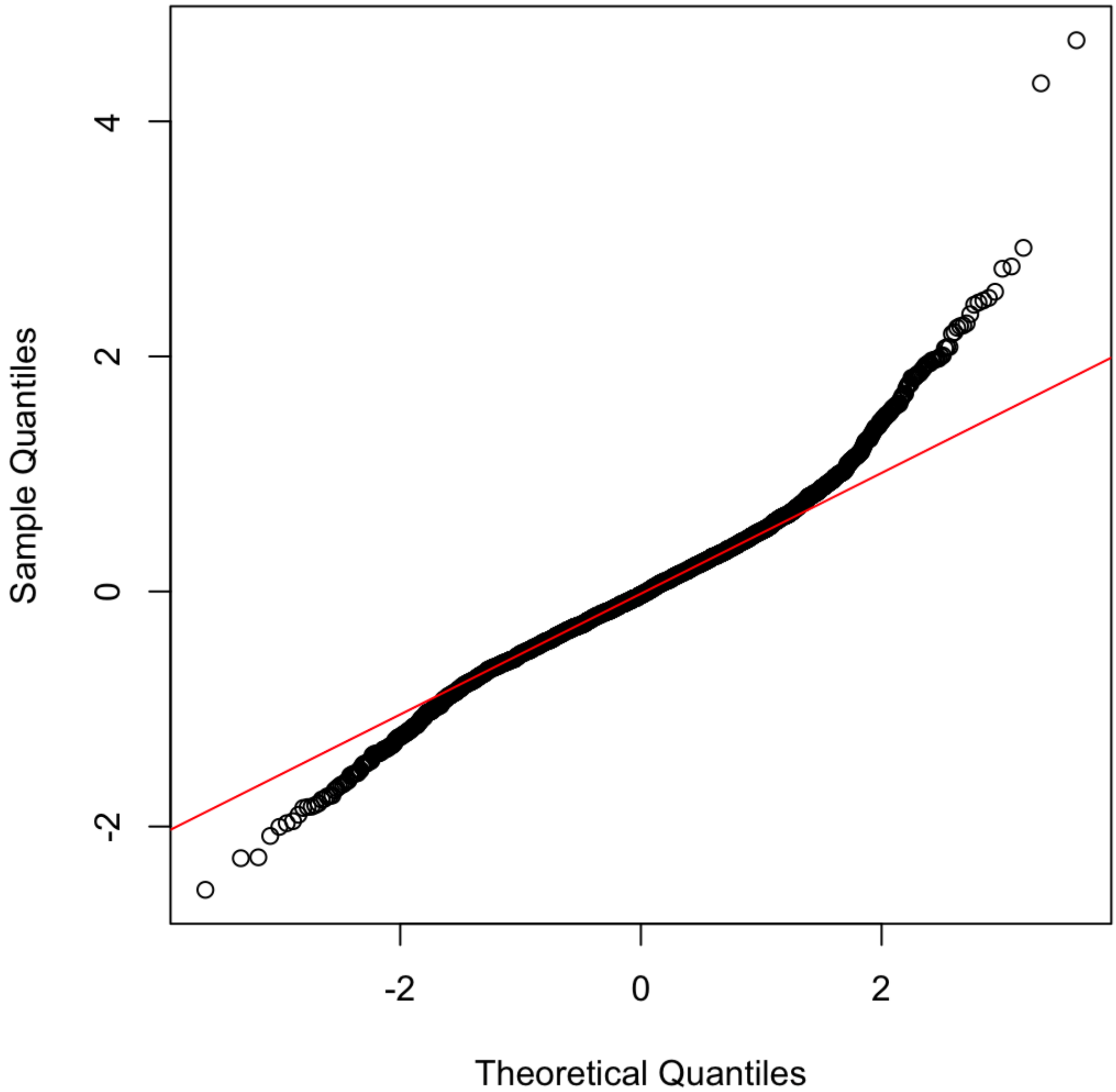


Table B4. Study 1: Descriptive Statistics

Variables	Mean	SD	Median 25 %ile	Median 50%ile	Median 75%ile	Skewness	Kurtosis
Emotionality flips	0.53	0.12	0.5	0.56	0.6	−1.72	4.15
Days Elapsed	789.93	543.22	357	668	1283	0.41	−1.04
Number of sentences	350.06	538.02	27	93	451	2.16	4.60
Average emotionality	0.15	0.16	0.07	0.17	0.24	−0.73	4.00
Fluency	−1250.33	2068.30	−1512.10	−715.80	−183.80	−5.66	52.57
Duration	2696.33	3632.77	379	1028	3727	2.56	10.78
Number of subscribers	3463012.5	3384141.69	1360000	2550000	4740000	1.97	4.14

Table B5. Study 1: Correlation Coefficients

	Correlations					
Variables	1	2	3	4	5	6
1. Emotionality flips						
2. Days Elapsed	0.111^{***}					
3. Number of sentences	0.128^{***}	-0.053^{***}				
4. Average emotionality	0.095^{***}	0.061^{***}	0.128^{***}			
5. Fluency	-0.229^{***}	0.016^{**}	-0.212^{***}	-0.075^{***}		
6. Duration	0.256^{***}	-0.047^{***}	0.799^{***}	0.084^{***}	-0.293^{***}	
7. Number of subscribers	-0.147^{***}	0.059^{***}	0.170^{***}	0.022^{***}	0.170^{***}	-0.092^{***}
Note:	$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$					

Table B6. Study 2: Descriptive Statistics

Variables	Mean	SD	Median 25 %ile	Median 50%ile	Median 75%ile	Skewness	Kurtosis
Emotionality flips	0.48	0.09	0.42	0.48	0.53	−0.83	3.43
Days Elapsed	2251.54	1407.97	956	2232	3392	0.15	−1.15
Number of sentences	106.25	54.94	67	100	137	0.87	1.45
Average emotionality	0.75	0.63	0.99	0.99	0.99	−2.30	3.42
Fluency	70.21	33.40	66.47	71.85	76.66	−34.42	1470.63
Duration	788.63	310.32	575	798	1006	0.30	0.76
Number of subtitles	26.06	8.81	21	26	31	0.08	1.44

Table B7. Study 2: Correlation Coefficients

	Correlations					
Variables	1	2	3	4	5	6
1. Emotionality flips						
2. Days Elapsed	-0.197^{***}					
3. Number of sentences	-0.076^{***}	0.323^{***}				
4. Average emotionality	-0.079^{***}	0.088^{***}	0.069^{***}			
5. Fluency	0.139^{***}	0.044^{*}	0.162^{***}	0.056^{**}		
6. Duration	0.084^{***}	0.226^{***}	0.805^{***}	0.002	0.39^{*}	
7. Number of subtitles	-0.095^{***}	0.460^{***}	-0.100^{***}	0.018	−0.018	-0.202^{***}
Note:	$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$					

Web Appendix C: Results for Individual Components of the Dependent Variable

Study 1: Robustness to Components of Audience Engagement as DVs

The number of views (skewness = 10.68), the number of likes (skewness = 12.81), and the number of comments (skewness = 15.88) were skewed. Therefore, we log-transform these three dependent measures, consistent with the log-transformation of the main DV (the sum of the three counts).

Table C1. Study 1: Regression estimates for different components of audience engagement

DV:	log (Number of views)	log (Number of comments)	log (Number of likes)
	(I)	(II)	(III)
Emotionality flips	0.928^{***}(0.341)	1.893^{***} (0.374)	0.554^{*} (0.325)
Days Elapsed	0.002^{*} (0.001)	0.001 (0.001)	0.001 (0.001)
Number of sentences	0.0004^{**} (0.0001)	0.001^{***} (0.0000)	0.003^{*} (0.0000)
Average emotionality	-0.664^{***}(0.133)	-0.768^{***} (0.140)	-0.570^{***} (0.128)
Fluency	-0.00004^{***} (0.0000)	-0.0001^{***} (0.0000)	-0.00004^{***} (0.0000)
Duration	0.0001^{***} (0.0000)	0.00003^{***} (0.0000)	0.00003^{*} (0.0000)
Number of subscribers	−0.0000 (0.0000)	0.0000^{**} (0.0000)	0.0000 (0.0000)
Copula term for IV	0.003 (0.011)	-0.019^{*} (0.010)	0.015 (0.010)
25 LDA Topic weights	✓	✓	✓
Channel FEs	✓	✓	✓
Year FEs	✓	✓	✓
Month FEs	✓	✓	✓
Day FEs	✓	✓	✓
Constant	8.359^{***} (1.431)	1.662 (1.981)	3.670^{*} (1.878)
Observations	33,584	33,584	33,584
Log Likelihood	−52,755.570	−54,391.670	−51,896.890
AIC	105,719.100	108,991.300	104,001.800
Note:			$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Study 2: Robustness to Components of Audience Engagement as DVs

The number of views (skewness = 8.63) and the number of comments (skewness = 8.84) were skewed. Therefore, we log-transform these two dependent measures, consistent with the log-transformation of the main DV (the sum of the two counts).

Table C2. Study 2: Regression estimates for different components of audience engagement

DV:	**log (Number of views)**	**log (Number of comments)**
	(I)	(II)
Emotionality flips	0.674^{***}(0.173)	0.865^{***} (0.187)
Days Elapsed	0.003^{**} (0.001)	0.002 (0.001)
Number of sentences	0.003^{***} (0.0004)	0.002^{***} (0.0004)
Average emotionality	0.031^{*}(0.017)	-0.073^{***} (0.019)
Fluency	0.0002 (0.0003)	0.0002 (0.0004)
Duration	0.0004^{***} (0.0001)	0.001^{***} (0.0000)
Number of subtitles	-0.063^{***} (0.002)	0.058^{***} (0.002)
Copula term for IV	0.018 (0.017)	0.004 (0.018)
20 LDA Topic weights	✓	✓
Gender FEs	✓	✓
Year FEs	✓	✓
Month FEs	✓	✓
Day FEs	✓	✓
Constant	-6.112^{***} (6.810)	11.670 (7.460)
Observations	3,381	3,381
Log Likelihood	−3,171.966	−3,481.522
AIC	6,467.933	7,087.044
Note:		$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Web Appendix D: Flesch Reading Ease Score

Flesch Reading Ease Score = 206.835 – (1.015 × ASL) – (84.6 × ASW), where ASL is the average sentence length (number of words divided by number of sentences), and ASW is the average number of syllables per word (number of syllables divided by number of words).

Web Appendix E: Topic Modeling Estimation

Prior research (Berger et al. 2021; Packard et al. 2023) has reported that audience engagement with content is a function of the topics it covers. Therefore, we built a topic model—specifically, a latent Dirichlet allocation (LDA; Blei, 2012)—on the corpus of the titles of the 33,598 episodes. Following prior research (Packard et al. 2023), our model utilized Gibbs sampling over 5,000 iterations with a random seed as the starting point. We began with 10 topics and increased the number of topics in increments of 5 until no further improvement (reduction) in perplexity was observed (Blei 2012; Chang et al. 2009). The model with 25 topics produced the lowest perplexity score.

Web Appendix F: Robustness Checks for Study 1

Robustness to Random-Effects Regression

We test our findings' robustness to a random-effects (rather than a fixed-effects) specification for three reasons. First, the podcast episodes are nested within channels and potentially within periods. Therefore, YouTube may recommend related episodes from another season or period to the audience. One might reason that the differences between channels represent random variability, which requires the inclusion of random effects to account for this variability instead of clustering error terms within channels (Gelman and Hill 2006). Including random effects helps account for variability within and between these groups, potentially leading to a more accurate and reliable model (Bates et al. 2015; Gelman and Hill 2006). Second, a random-effects regression accounts for the hierarchical nature of the data, ensuring more precise parameter estimates and reducing the risk of bias (Harrison et al. 2018). Third, one might reason that episodes belonged to different channels (n = 50), so we also need to include a random effect for the channel to account for unobserved variables such as the podcaster's idiosyncratic style. Therefore, we control for the random effect of each episode's channel.[14] The independent variable, dependent variable, covariates, and estimation method are similar to the fixed-effects specification.

Results. The results remain robust under the random-effects specification (Table F1, Column I: $b_{\textit{Emotionality flips} \rightarrow \textit{Audience engagement}}$ = 0.947, p < .001, Exp(b) = 2.58). The interpretation is that a 1% increase in the emotionality flips in a podcast episode is associated with a 158%

[14] A Hausman test ($X^2(4)$ = 153.38; p < 0.001) indicated that fixed-effect specification is more appropriate when channel, year, month, and day are used as covariates. Therefore, fixed-effect specifications are used in the main model for these variables. When Hausman tests were conducted separately, the results confirmed our decision to choose fixed-effects specifications in the main model again. Results showed that the channel ($X^2(1)$ = 185.53; p < 0.001), year ($X^2(1)$ = 36.16; p < 0.001), month ($X^2(1)$ = 23.72; p < 0.001), and day ($X^2(1)$ = 5.82; p < 0.05) are not correlated with the regressor.

increase in audience engagement with the episode.[15] This effect size is similar to 157% in our fixed-effects specification, suggesting that the observed effect is unlikely due to specification.

Robustness to an Alternative Rule-based Method of Measuring Emotionality

We test whether our results hold when emotionality is measured using a software program other than VADER. Prior research has used TextBlob as an alternative to VADER (Hwang and Lee 2024; Mousavi and Gu 2023; Rivera et al. 2023; Wang et al. 2023). TextBlob measures the sentiment of sentences on a scale from −1 to +1, where values closer to −1 indicate negative emotionality and values closer to +1 indicate positive emotionality. The methods for calculating the independent variable, dependent variable, and covariates, and for estimation, are identical to those in our main analysis.

Results. We continue to find a positive association between emotionality flips in a podcast episode and audience engagement with the episode (Table F1, Column II: $b_{Emotionality\ flips \rightarrow Audience\ engagement}$ = 0.824, $p < .01$, *Exp*(b) = 2.28).[16] The interpretation is that a 1% increase in emotionality flips in a podcast episode is associated with a 128% increase in audience engagement with the episode.[17] Again, the effect size of 128% is close to the 157% we obtained from VADER's measure, suggesting that our measured effect is unlikely to be a function of the software program used to measure emotionality.

Robustness to an Alternative Deep Learning–Based Method of Measuring Emotionality

While our main analysis uses VADER, a rule-based and lexicon-driven approach, we complement it with a modern, contextual language model to capture nuanced emotional

[15] $(Exp\ (b) - 1)\ \times\ 100 = (Exp\ (0.947) - 1)\ \times\ 100 = 158\%$.
[16] Shapiro-Wilk ($W = 0.94$, $p < .001$), Anderson-Darling (A = 64.09, $p < .001$), Kolmogorov-Smirnov (D = 0.08, $p < .001$), and Cramer-von Mises tests of normality ($W = 9.63$, $p < .001$), when TextBlob is used to measure emotionality flips, reject the null hypotheses of normality of emotionality flips and thus support our inclusion of the Gaussian copula term to control for the endogeneity of emotionality flips.
[17] $(Exp\ (b) - 1)\ \times\ 100 = (Exp\ (0.824) - 1)\ \times\ 100 = 128\%$.

tone. Specifically, we trained a DistilBERT-based sentiment regression model to output continuous sentiment scores at the sentence level.

Models like BERT, RoBERTa, and DistilBERT are typically trained as classifiers (outputting class probabilities rather than a sentiment magnitude). However, we prefer a continuous measure of sentiment. We trained a DistilBERT-based sentiment regression model using a modified version of the SST-2 dataset, which originally contains binary sentiment labels (0 for negative, 1 for positive). We transformed these labels into a pseudo-continuous format, allowing the model to be fine-tuned as a regressor rather than a classifier. The regression head consists of a single linear layer followed by an activation function, enabling the model to produce sentiment scores in the range of [−1, +1]. Although the original labels are binary, the model learns to infer degrees of sentiment strength from sentence-level context, capturing subtle differences in sentiment even without explicit intensity labels. This setup enables us to obtain real-valued sentiment scores that reflect both polarity and confidence. Next, we analyzed emotionality flips' effect on audience engagement. The methods for calculating the independent and dependent variables, the covariates, and the estimation are the same as in our main analysis.

Results. Emotionality flips continues to be positively associated with audience engagement with podcasts (Table F1, Column III: $b_{\text{Emotionality flips} \rightarrow \text{Audience engagement}} = 0.933$, $p < .01$ *Exp* (b) = 2.54).[18] The interpretation is that a 1% increase in the emotionality flips in a podcast episode is associated with a 154% increase in engagement with that episode.[19] The

[18] Shapiro-Wilk ($W = 0.73$, $p < .001$), Anderson-Darling (A = Inf, $p < .001$), Kolmogorov-Smirnov (D = 0.20, $p < .001$), and Cramer-von Mises tests of normality ($W = 16.74$, $p < .001$), when DistilBERT is used to measure emotionality flips, reject the null hypotheses of normality of emotionality flips and thus support our inclusion of the Gaussian copula term to control for the endogeneity of emotionality flips.

[19] $(Exp\ (b) - 1) \times 100 = (Exp\ (0.933) - 1) \times 100 = 154\%$.

effect size of 154% is close to the 157% we obtained from VADER's measure, providing further confidence in our estimate.

Robustness to ChatGPT-Based Measurement of Emotionality

Next, we employed ChatGPT (GPT-4o) to generate emotionality scores based on natural language understanding. GPT-4o has demonstrated near-human performance in interpreting nuanced language and affective meaning (Bojic et al. 2023), making it a compelling alternative for assessing emotional dynamics in narrative content. We randomly selected 10% of our podcast episodes. Next, we prompted ChatGPT to rate each sentence's emotionality on a standardized scale from −1 to +1, where −1 indicates very negative emotionality, and +1 indicates very positive emotionality (see below for the full prompt and procedure). Using these sentence-level scores, we computed emotionality flips following the same method as in our main analysis. The methods for calculating the independent and dependent variables, the covariates, and the estimation are the same as in our main analysis.

Results. We continue to find a positive association between the emotionality flips in a podcast episode and audience engagement with the episode (Table F1, Column IV: $b_{\textit{Emotionality flips} \rightarrow \textit{Audience engagement}}$ = 0.357, p < .01, *Exp* (b) = 1.43).[20] The interpretation is that a 1% increase in emotionality flips in a podcast episode is associated with a 43% increase in engagement with the episode.[21] The effect size drops from 157% (using the VADER measure) to 43% (using the ChatGPT measure) for three reasons. First, GPT-4o tends to produce more moderate sentiment scores, which reduces the number of detected emotionality flips

[20] Shapiro-Wilk (W = 0.85, p < .001), Anderson-Darling (A = 6.16, p < .001), Kolmogorov-Smirnov (D = 0.18, p < .001), and Cramer-von Mises tests of normality (W = 0.8, p < .05), when two sentences are used as the chunk size, reject the null hypotheses of normality of emotionality flips and thus support our inclusion of the Gaussian copula term to control for the endogeneity of emotionality flips.

[21] $(Exp\ (b) - 1)\ \times\ 100 = (Exp\ (0.357) - 1)\ \times\ 100 = 43\%$.

(Vamvourellis and Mehta 2025). Second, we used only 10% of the dataset, which lowers statistical power and increases variability in estimates. Third, relying on prompts introduces subjectivity and variation, which can weaken the measured relationship (Cuellar et al. 2025). Notwithstanding these three reasons, the estimated coefficient remains positive and significant, reinforcing our main effect's robustness.

Robustness to an Alternative Unit(s) of Measurement of Emotionality[22]

One might reason that a sentence is not the most appropriate unit for measuring emotionality flips, since some sentences describe the previous one, introduce the next, or serve as a transition. Therefore, we measure the emotionality score (VADER's *compound* variable) in units of *two* sentences. The methods for calculating the independent and dependent variables, the covariates, and the estimation are the same as in our main analysis.

Results. Mirroring previous findings, we find a positive association between the emotionality flips in a podcast episode and engagement with the episode (Table F1, Column V: $b_{Emotionality\ flips\ \rightarrow\ Audience\ engagement} = 0.898$, $p < .01$, $Exp(b) = 2.45$).[23] The interpretation is that a 1% increase in the emotionality flips in a podcast episode is associated with a 145% increase in audience engagement.[24] Again, the effect size remains similar to that in our main specification.

---

[22] See Table F1 (columns X–XII) for results using three-, four-, and five-sentence chunk sizes. The results remained the same, demonstrating that the effect of emotionality flips on engagement is robust to alternative chunking specifications.

[23] Shapiro-Wilk ($W = 0.75$, $p < .001$), Anderson-Darling (A = 424.14, $p < .001$), Kolmogorov-Smirnov (D = 0.20, $p < .001$), and Cramer-von Mises tests of normality ($W = 75.09$, $p < .001$), when two sentences are used as the chunk size, reject the null hypotheses of normality of emotionality flips and thus support our inclusion of the Gaussian copula term to control for the endogeneity of emotionality flips.

[24] $(Exp\ (b) - 1) \times 100 = (Exp\ (0.898) - 1) \times 100 = 145\%$.

Robustness to Alternative Measures of the Dependent Variable

YouTube algorithms suggest podcast episodes to audiences whose viewing habits indicate sustained interest in similar content. Moreover, 1,409 episodes in our dataset were released within a month, suggesting a consistent output schedule that likely contributed to a steady rise in engagement. Therefore, instead of using the number of days elapsed (between the episode posting date and our data collection date of April 29, 2024) as a control variable, we divide our DV (the sum of the number of views, likes, and comments) by the number of days elapsed. The skewness of the resulting ratio was 16.76, and therefore, we natural log-transformed the ratio values.

Results. We continue to find a positive association between the emotionality flips in a podcast episode and engagement with the episode (Table F1, Column VI: $b_{Emotionality\ flips \rightarrow Audience\ engagement}$ = 0.936, $p < .001$, Exp(b) = 2.55). The interpretation is that a 1% increase in the emotionality flips in a podcast episode is associated with a 155% increase in audience engagement.[25] Again, the effect size remains similar to that in our main specification.

Robustness to Controlling for Sentiment Volatility

Berger et al. (2021) introduced a measure that explains user engagement with content. They define "sentiment volatility as the standard deviation (SD) of differences in sentiment between consecutive chunks of an experience" (Berger et al., 2021, p. 237). We reason that our emotionality flips variable is conceptually distinct from Berger et al.'s (2021) sentiment volatility in three ways. First, flips capture changes in the direction of the sentiment slope, regardless of their magnitude or extremity. In contrast, sentiment volatility captures the extremity and magnitude of differences between chunk sizes. Second, flips represent the number of times emotionality changes direction from one unit (e.g.,

[25] $(Exp\ (b) - 1)\ \times\ 100 = (Exp\ (0.936) - 1)\ \times\ 100 = 155\%.$

sentiment) to another throughout the entire narrative content. In contrast, volatility is measured between consecutive chunks, assuming a user's consumption unit spans multiple chunks. Third, flips use a predefined chunk size (e.g., sentences), whereas volatility does not rely on a predefined chunk size. Notwithstanding these distinctions, we test whether emotionality flips' effect on engagement stays after our regression includes sentiment volatility. Following Berger et al., (2021), we measure an episode's *Sentiment volatility* as the standard deviation (SD) of the differences between consecutive sentences' *compound* values.[26]

Results. We estimate two regressions. First, we replace *Emotionality flips* with *Sentiment volatility* in our regression and find that *Sentiment volatility* is positively associated with *Audience engagement* (untabulated results: $b_{\text{Sentiment volatility} \rightarrow \text{Audience engagement}}$ = 0.400, $p < .01$). Thus, we replicate Berger et al.'s (2021) findings. Second, we include *Emotionality flips* in this regression. *Sentiment volatility*'s beta is insignificant (Table F1, Column VII: $b_{\text{Sentiment volatility} \rightarrow \text{Audience engagement}}$ = 0.088, $p = ns$), whereas *Emotionality flips* is positively associated with audience engagement (Table F1, Column VII: $b_{\text{Emotionality flips} \rightarrow \text{Audience engagement}}$ = 0.911, $p < .001$, *Exp*(b) = 2.49). The theoretical insight is that *Emotionality flips* explains audience engagement even after we control for sentiment volatility. The interpretation is that a 1% increase in the emotionality flips in a podcast episode is associated with a 149% increase in audience engagement.[27] Again, the effect size remains similar to that in our main specification.

[26] Consider an episode with three sentences. Sentence 1's compound value is 0.5, and sentence 2's value is 0.4. So, the difference is 0.1. Next, sentence 3's compound value is -0.1, leading to a difference of 0.5 from the previous sentence's compound value. So, we calculate SD of 0.1 and 0.5.

[27] $(Exp\ (b) - 1) \times 100 = (Exp\ (0.911) - 1) \times 100 = 149\%$.

Robustness to Controlling for the Magnitude of Emotionality Flips

Emotionality flips captures the number of times sentence-level emotionality changes directions (or flips). It does *not* account for the magnitude of the flips. Therefore, we control for the emotionality changes between consecutive sentences, averaged over the episode's sentences.[28]

Results. Once again, *Emotionality flips* is positively associated with audience engagement (Table F1, Column VIII: $b_{\textit{Emotionality flips} \rightarrow \textit{Audience engagement}}$ = 0.980, $p < .001$, *Exp*(b) = 2.66). In contrast, the mean steepness of changes in emotionality slope does not statistically significantly associate with engagement (Table F1, Column VIII: $b_{\textit{Mean steepness} \rightarrow \textit{Audience engagement}}$ = −0.171, $p = ns$). The theoretical insight is that the number of emotionality flips, rather than their magnitude, drives engagement. The interpretation is that a 1% increase in the emotionality flips in a podcast episode is associated with a 166% increase in audience engagement.[29] Again, the effect size remains similar to that in our main specification.

Robustness to Controlling for Semantic Progression

We control for semantic progression to check whether emotionality flips' effect on audience engagement is confounded by the narrative's broader semantic characteristics. Our revised regression includes Toubia et al.'s (2021) three measures, which capture the evolution of meaning across discourse using high-dimensional embeddings from natural language processing models. Specifically, (1) *Semantic speed* captures the average semantic distance between adjacent chunks (i.e., the rate at which the narrative progresses through

---

[28] Consider an episode with three sentences and compound values of [0.5, 0.4, −0.1]. First, we calculate the absolute differences between each consecutive pair: (|0.4 - 0.5| = 0.1) and (|−0.1 − 0.4| = 0.5). The absolute difference measures the magnitude of emotionality slope between consecutive sentences. Second, we sum these differences (0.1 + 0.5 = 0.6), thus computing the total emotionality slope in the episode. Third, we divide the sum by the number of intervals (i.e., 2), resulting in 0.3. The final score is the average emotionality slope in the episode.

[29] $(Exp\ (b) - 1)\ \times\ 100 = (Exp\ (0.980) - 1)\ \times\ 100 = 166\%.$

semantic space), (2) *Semantic volume* quantifies the overall area covered (i.e., how much ground the narrative spans), and (3) *Semantic circuitousness* measures how directly the narrative moves from start to finish (i.e., whether the semantic path is linear or winding).

Whereas emotionality flips operate in the affective domain, tracking directional changes in emotional valence within a narrative, semantic progression operates in the cognitive domain, capturing shifts in meaning and topic structure. Emotionality flips are calculated at the sentence level, capturing fine-grained fluctuations in emotional tone, while semantic progression is computed over larger text chunks (e.g., 250-word blocks), reflecting broader thematic movements. The two constructs also differ theoretically: emotionality flips are grounded in research on emotional engagement and narrative transportation, whereas semantic progression draws from cognitive and linguistic theories of idea development and discourse structure. Together, these distinctions support both constructs' inclusion in our model to isolate emotionality flips' power to explain variation in audience engagement.

Results. We estimate two regressions. First, we replace *Emotionality flips* with *Semantic progression* variables in our regression. Results showed that faster-paced podcasts are associated with lower engagement, while those that cover more ground and follow more circuitous paths receive more engagement (untabulated results: $b_{Speed \rightarrow \text{Audience engagement}} = -0.132$, $p < .001$; $b_{Volume \rightarrow \text{Audience engagement}} = 0.164$, $p < .001$; $b_{Circuitousness \rightarrow \text{Audience engagement}} = 0.159$, $p < .001$). These results mirror Toubia et al.'s (2021) findings, suggesting that podcasts—like academic articles—reward intellectual exploration and conceptual richness, with audiences favoring depth over speed. Second, we include *Emotionality flips* in this regression. Results confirmed again that *Emotionality flips* is positively associated with audience engagement (Table F1, Column IX: $b_{\text{Emotionality flips} \rightarrow \text{Audience engagement}} = 0.408$, $p < .05$, *Exp*(b) = 1.50). The theoretical insight is that *Emotionality flips* is a robust explainer of audience engagement.

Study 1 Robustness Checks

Table F1 (Columns I to III). Robustness checks for Study 1

	DV = log (Audience engagement)		
	(I) Random effects	(II) TextBlob	(III) DistilBERT
Emotionality flips	0.947*** (0.074)	0.824** (0.341)	0.933** (0.405)
Fluency	−0.00004*** (0.0000)	−0.00004*** (0.0000)	−0.0004*** (0.0001)
Number of sentences	0.0004*** (0.00003)	0.0004*** (0.00003)	0.0003** (0.0002)
Average emotionality	−0.661*** (0.133)	−0.653*** (0.133)	−0.622*** (0.127)
Days elapsed	0.002** 0.001)	0.002* (0.001)	0.002 (0.001)
Duration	0.0001*** (0.0000)	0.0001*** (0.0000)	0.0001*** (0.0002)
Number of subscribers	−0.0000 (0.0000)	−0.0000*** (0.0000)	−0.0000 (0.0000)
Copula term for IV	−0.003 (0.010)	−0.003 (0.010)	0.001 (0.010)
25 LDA Topic weights	✓	✓	✓
Channel FEs		✓	✓
Channel REs	✓		
Year FEs	✓	✓	✓
Month FEs	✓	✓	✓
Day FEs	✓	✓	✓
Constant	8.023*** (1.426)	7.911*** (1.788)	7.961*** (1.811)
Observations	33,584	33,584	33,584
Log likelihood	−53,018.980	−52,691.170	−52,692.350
AIC	106,154.000	105,590.300	105,592.700
BIC	106,642.400		
Note:			* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table F1 (continued; Columns IV to VI). Robustness checks for Study 1

	DV = log (Audience engagement)		DV = log (Audience engagement ÷ # days elapsed)
	(IV) ChatGPT	(V) Chunks of two sentences	(VI) Alternative measure of DV
Emotionality flips	0.357** (0.142)	0.898** (0.353)	0.936*** (0.332)
Fluency	−0.00005***(0.0000)	−0.00004***0.0000)	−0.00004*** (0.0000)
Number of sentences	0.0004*** (0.0001)	0.0003** (0.00003)	0.0004** (0.00003)
Average emotionality	−0.495*** (0.148)	−0.639*** (0.128)	−0.651*** (0.131)
Days elapsed	0.005** (0.002)	0.002* (0.001)	
Duration	0.0001***(0.0000)	0.0001*** (0.0000)	0.0001*** (0.0000)
Number of subscribers	−0.0000** (0.0000)	−0.0000 (0.0000)	−0.0000 (0.0000)
Copula term for IV	−0.022 (0.028)	0.011 (0.008)	−0.003 (0.009)
25 LDA Topic weights	✓	✓	✓
Channel FEs	✓	✓	✓
Year FEs	✓	✓	✓
Month FEs	✓	✓	✓
Day FEs	✓	✓	✓
Constant	1.990*** (4.641)	7.848*** (1.777)	3.150*** (0.492)
Observations	3,360	33,584	33,584
Log likelihood	−5,250.421	−52,597.710	−53,438.970
AIC	10,708.840	105,403.400	107,083.900
Note:			$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table F1 (continued; Columns VII to X). Robustness checks for Study 1

	DV = log (Audience engagement)		
	(VII) Including sentiment volatility	(VIII) Including mean steepness	(IX) Including semantic progression
Emotionality flips	0.911*** (0.344)	0.980*** (0.362)	0.408* (0.235)
Fluency	−0.00004*** (0.0000)	−0.00004*** (0.0000)	−0.00003*** (0.0000)
Number of sentences	0.0004** (0.00003)	0.0004** (0.00003)	0.0001 (0.0001)
Average emotionality	−0.645*** (0.133)	−0.686*** (0.134)	−0.632*** (0.118)
Days elapsed	0.002* (0.001)	0.002* (0.001)	0.002** (0.001)
Duration	0.0001*** (0.0000)	0.0001*** (0.0000)	0.00004** (0.0000)
Number of subscribers	−0.0000 (0.0000)	−0.0000 (0.0000)	−0.0000 (0.0000)
Sentiment volatility	0.088 (0.138)		
Mean steepness		−0.171 (0.178)	
Average speed			−0.119*** (0.038)
Normalized volume			0.154*** (0.053)
Circuitousness			0.154*** (0.040)
Narrativity			
Copula term for IV	−0.006 (0.009)	0.002 (0.011)	0.005 (0.009)
25 LDA Topic weights	✓	✓	✓
Channel FEs	✓	✓	✓
Year FEs	✓	✓	✓
Month FEs	✓	✓	✓
Day FEs	✓	✓	✓
Constant	7.810*** (1.856)	7.954*** (1.851)	7.676*** (1.713)
Observations	33,584	33,584	30,754
Log likelihood	−52,676.290	−52,673.300	−46,053.260
AIC	105,562.600	105,556.600	92,320.520
Note:	$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$		

Table F1 (continued; Columns XI to XIII). Robustness checks for Study 1

	DV = Audience Engagement		
	(X) Chunks of three sentences	(XI) Chunks of four sentences	(XII) Chunks of five sentences
Emotionality flips	0.862** (0.345)	0.799** (0.337)	0.793** (0.332)
Fluency	−0.00003*** (0.00001)	−0.00003*** (0.00001)	−0.00003*** (0.00001)
Number of sentences	0.003* (0.0002)	0.003* (0.0002)	0.003* (0.0002)
Average emotionality	−0.634*** (0.123)	−0.623*** (0.124)	−0.630*** (0.124)
Days elapsed	0.002* (0.001)	0.002** (0.001)	0.002** (0.001)
Duration	0.0001*** (0.0000)	0.001*** (0.00002)	0.001*** (0.00002)
Number of subscribers	−0.0000 (0.0000)	−0.0000 (0.0000)	−0.0000 (0.0000)
Copula term for IV	−0.007 (0.010)	−0.011 (0.011)	−0.022* (0.012)
25 LDA Topic weights	✓	✓	✓
Channel FEs	✓	✓	✓
Year FEs	✓	✓	✓
Month FEs	✓	✓	✓
Day FEs	✓	✓	✓
Constant	7.935*** (1.762)	7.878*** (1.737)	7.632*** (1.721)
Observations	33,369	33,369	33,369
Log likelihood	−52,162.000	−52,170.590	−52,168.090
AIC	104,532.000	104,549.200	104,544.200
Note:	* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$		

Study 2 Robustness Checks

Table F2 (Columns I to III). Robustness checks for Study 2

	DV = log (Audience engagement)		
	(I) TextBlob	(II) DistilBERT	(III) ChatGPT
Emotionality flips	0.548*** (0.169)	0.418* (0.216)	0.879*** (0.189)
Fluency	0.0003 (0.0003)	0.004 (0.0004)	0.0001 (0.0003)
Number of sentences	0.003*** (0.0004)	0.0003*** (0.001)	0.003*** (0.0004)
Average emotionality	0.015 (0.133)	0.023 (0.017)	−0.495*** (0.148)
Days elapsed	0.003** (0.001)	0.003** (0.001)	0.003*** (0.001)
Duration	0.0004*** (0.0001)	0.005*** (0.0001)	0.0004*** (0.0001)
Number of subtitles	0.064*** (0.002)	0.064*** (0.002)	0.064*** (0.002)
Copula term for IV	0.018 (0.016)	−0.013 (0.016)	−0.019 (0.016)
20 LDA Topic weights	✓	✓	✓
Gender FEs	✓	✓	✓
Year FEs	✓	✓	✓
Month FEs	✓	✓	✓
Day FEs	✓	✓	✓
Constant	−5.884 (6.815)	−5.785 (6.842)	−5.889 (6.817)
Observations	3,381	3,381	3,381
Log likelihood	−3,177.323	−3,182.302	−3,178.063
AIC	6,478.646	6,488.604	6,480.125
Note:	$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$		

Table F2 (Columns IV to VI). Robustness checks for Study 2

	DV = log (Audience engagement)	DV = log (Audience engagement ÷ # days elapsed)	DV = log (Audience engagement)
	(IV) Chunks of two sentences	(V) Alternative measure of DV	(VI) Including sentiment volatility
Emotionality flips	0.415*** (0.094)	0.661*** (0.175)	0.618*** (0.209)
Fluency	0.0003 (0.0003)	0.0002 (0.0003)	0.003* (0.001)
Number of sentences	0.003*** (0.0004)	0.003*** (0.0004)	0.003*** (0.0004)
Average emotionality	0.032* (0.017)	0.032* (0.018)	0.038** (0.018)
Days elapsed	0.003** (0.001)		0.002* (0.001)
Duration	0.0004*** (0.0001)	0.0004*** (0.0001)	0.0004*** (0.0001)
Number of subtitles	0.063*** (0.002)	0.059*** (0.002)	0.063*** (0.002)
Sentiment volatility			0.289 (0.180)
Mean steepness			
Average speed			
Normalized volume			
Circuitousness			
Copula term for IV	−0.005 (0.016)	0.026 (0.009)	0.005 (0.017)
20 LDA Topic weights	✓	✓	✓
Gender FEs	✓	✓	✓
Year FEs	✓	✓	✓
Month FEs	✓	✓	✓
Day FEs	✓	✓	✓
Constant	5.418 (6.813)	0.396 (1.232)	−6.276*** (6.829)
Observations	3,381	3,381	3,381
Log likelihood	−3,175.908	−3,263.025	−3,161.518
AIC	6,475.815	6,648.051	6,649.036
Note:	$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$		

Table F2 (Columns VII to VIII). Robustness checks for Study 2

	DV = log (Audience engagement)	
	(VII) Including mean steepness	(VIII) Including semantic progression
Emotionality flips	0.790*** (0.210)	0.626*** (0.186)
Fluency	0.003** (0.001)	−0.00003*** (0.0000)
Number of sentences	0.003*** (0.0004)	0.003*** (0.0004)
Average emotionality	0.037** (0.018)	0.026 (0.017)
Days elapsed	0.003** (0.001)	0.003** (0.001)
Duration	0.0004*** (0.0001)	0.0003*** (0.0001)
Number of subtitles	0.063*** (0.002)	0.064*** (0.002)
Sentiment volatility		
Mean steepness	0.308 (0.197)	
Average speed		−0.057*** (0.019)
Normalized volume		0.072*** (0.020)
Circuitousness		0.012 (0.019)
Narrativity		
Copula term for IV	−0.020 (0.017)	0.003 (0.017)
20 LDA Topic weights	✓	✓
Gender FEs	✓	✓
Year FEs	✓	✓
Month FEs	✓	✓
Day FEs	✓	✓
Constant	−6.794*** (6.832)	−6.363 (6.831)
Observations	3,381	3,358
Log likelihood	−3,160.909	−3,140.423
AIC	6,447.818	6,410.847
Note:	$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$	

Table F2 (Columns X to XII). Robustness checks for Study 2

	DV = Audience Engagement		
	(IX) Chunks of three sentences	(X) Chunks of four sentences	(XI) Chunks of five sentences
Emotionality flips	0.443*** (0.148)	0.510*** (0.131)	0.309** (0.124)
Fluency	0.0004 (0.0003)	0.0004 (0.0003)	0.0004 (0.0003)
Number of sentences	0.003*** (0.00004)	0.003*** (0.00004)	0.002*** (0.00004)
Average emotionality	0.021 (0.017)	0.023 (0.017)	0.022 (0.017)
Days elapsed	0.003** (0.001)	0.003** (0.001)	0.003** (0.001)
Duration	0.0004*** (0.0000)	0.0004*** (0.0000)	0.0004*** (0.0000)
Number of subtitles	0.064*** (0.002)	0.063*** (0.002)	0.064*** (0.002)
Copula term for IV	−0.027 (0.017)	−0.041** (0.016)	0.012 (0.016)
20 LDA Topic weights	✓	✓	✓
Gender FEs	✓	✓	✓
Year FEs	✓	✓	✓
Month FEs	✓	✓	✓
Day FEs	✓	✓	✓
Constant	−5.784 (6.834)	−6.138 (6.832)	−6.158 (6.829)
Observations	3,381	3,381	3,381
Log likelihood	−3,186.539	−3,183.468	−3,183.806
AIC	6,497.079	6,490.937	6,491.611
Note:			$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

ChatGPT Prompt Used in Robustness Checks in Studies 1 and 2

{

Your task is to analyze a list of texts and generate a CSV file with two columns:

1. **ID**: A unique identifier for each text
2. **Score**: A normalized sentiment fluctuation score between 0 and 1

For each text, perform the following steps:

- Split the text into individual sentences.
- Assign a sentiment score to each sentence on a scale from -1 to +1, where:
  - -1 = very negative
  - 0 = neutral
  - +1 = very positive
- Store the sentence-level sentiment scores as a list (e.g., [-0.4, -0.1, 0.3, 0.2, 0.8]).

    - Ensure each sentence's score considers the sentiment of its neighboring sentences (previous and next) to reflect the magnitude of emotional change.
- Count the number of **emotionality flips**, defined as:
    - A change in the *direction* of sentiment movement — i.e., a switch from rising to falling sentiment, or vice versa.
    - For example, a transition from increasing positivity to decreasing positivity, or from decreasing negativity to increasing negativity.
- Each such change is considered a **peak** or **trough** in the sentiment trajectory.
- Compute the total number of peaks and troughs, and divide this by the total number of sentences to obtain a **normalized score**.
- Output the results in a CSV file with the specified columns.

}

Web Appendix G: Study 3 Stimuli

The objective of this study is to examine how audiences evaluate podcasts.

On the next page, you will watch a podcast. Please watch it carefully, as you will be asked questions about it.

<BREAK>

On the next page, please watch a podcast episode from YouTube.

Manipulation: 1 episode with low emotionality flips (low 10 percent) vs. 1 episode with high emotionality flips (Top 10 percent)

JockoPodcastOfficial:
Low: (emotionality extremes: 0.36)
https://www.youtube.com/watch?v=eZJl79lh2K8
High: (emotionality extremes: 0.66)
https://www.youtube.com/watch?v=bmYtP9uKwFU

<BREAK>

DV (1= Not at all, 7= Very much).

1. How interesting did you find this podcast episode?
2. How much did you like this podcast episode?

<BREAK>

Mediators:

Arousal (Berger 2011):

How does this episode make you feel? (1 to 7; “very low energy/very high energy,” “very passive/very active,” and “very mellow/very fired up”).

<BREAK>

Surprise:

To what extent do you think you were able to predict what will happen next in the story?

How unexpected were the turns of events in this study?

How easy was it for you to predict what would happen next at any given point in the story?

<BREAK>

Curiosity (1= Not at all, 7=Very much)

I am curious to listen to the next episode of this story.

I am eager to learn more about the narrator of the story.

I would like to know more about the main character of the story.

<BREAK>

Demographic questions

Web Appendix H: Study 4 Details

Low emotionality flips condition (Fluency = 68.36):

"Just before dawn in 1990, two men dressed as police knocked on the door of the Isabella Stewart Gardner Museum (score: 0.0). The guards calmly hesitated - it seemed routine - and opened the door (score: 0.0). Moments later, the calm was gone (score: 0.32). The guards were handcuffed, and thirteen masterpieces vanished into the dark (score: 0.54). For weeks, Boston buzzed with excitement and hope (score: 0.73). Maybe the art would return, maybe the thieves would make a mistake (score: −0.69). But time passed, and the hope painfully faded away (score: −0.19). The mystery still hangs in the museum, quiet but alive in every empty frame (score: 0.30)."

High emotionality flips condition (Fluency = 65.93):

"In 1990, two men disguised as police officers came to the Isabella Stewart Gardner Museum in Boston(score: −0.27). The guards thought it was a routine, harmless call, and let them in right away (score: 0.25). A few minutes later, the guards were tied up, and thirteen paintings were gone, taken quickly and without a sound (score: 0.0). The theft shocked the city (score: −0.32). Police searched across the country, from small towns to distant warehouses, but every clue led nowhere (score: 0.0). Police conducted an exhaustive investigation, but the paintings were not recovered (score: −0.06). The stolen works seemed to disappear, leaving only questions and haunting reminders of what was lost (score: −0.82). The mystery remains, steady and unchanged after all these years (score: 0.0)."

NFA Full Scale

Please indicate how much you agree or disagree with each statement (1 = Strongly Disagree; 7 = Strongly Agree):

1. I feel that I need to experience strong emotions regularly.

2. Emotions help people to get along in life.

3. I think that it is important to explore my feelings.

4. It is important need to be in touch with my feelings.

5. It is important for me to know how others are feeling.

6. If I reflect on my past, I see that I tend to be afraid of feeling emotions (R).

7. I find strong emotions overwhelming and therefore try to avoid them (R).

8. I would prefer not to experience either the lows or highs of emotion (R).

9. I do not know how to handle my emotions, so I avoid them (R).

10. Emotions are dangerous-they tend to get me into situations that I would rather avoid (R).

Web Appendix I: Study 5 Prompts and Podcasts

Table I1. Profile of participants defined for ChatGPT

	Characteristic Description		Characteristic Description
1	you are highly curious about science and technology	**26**	you dislike overly technical or procedural language
2	you are easily distracted while reading long texts	**27**	you enjoy suspense and emotional buildup in stories
3	you are deeply emotional and respond strongly to inspirational stories	**28**	you relate personally to stories of human bravery
4	you have little interest in historical topics	**29**	you have a short attention span
5	you are analytical and prefer factual details over emotions	**30**	you enjoy learning facts that few people know
6	you are impatient and lose focus when a story is too detailed	**31**	you are drawn to stories about innovation
7	you are fascinated by human achievement and perseverance	**32**	you dislike long factual lists or data points
8	you are skeptical about government programs and space missions	**33**	you like stories that make you feel awe or wonder
9	you enjoy vivid storytelling and sensory descriptions	**34**	you find it difficult to stay focused on non-fiction
10	you prefer concise summaries to long reports	**35**	you are fascinated by how people overcome obstacles
11	you are nostalgic and romanticize the past	**36**	you enjoy slow, reflective storytelling
12	you tend to overanalyze details	**37**	you dislike historical events unless they are emotional
13	you find technical descriptions difficult to visualize	**38**	you are methodical and focus on accuracy when judging content
14	you love learning about space exploration	**39**	you are inspired by real-world accomplishments
15	you often skim texts instead of reading carefully	**40**	you tend to prefer fiction over non-fiction
16	you are easily inspired by stories of courage	**41**	you appreciate stories about leadership and vision
17	you value emotional depth more than accuracy	**42**	you are indifferent to space exploration as a topic
18	you find science and engineering topics dull	**43**	you are attentive and patient while reading long stories
19	you are detail-oriented and appreciate precision	**44**	you get more engaged when the story includes emotion
20	you get bored quickly when the topic feels repetitive	**45**	you find procedural detail satisfying and credible
21	you like stories that emphasize teamwork and cooperation	**46**	you lose interest quickly when nothing dramatic happens
22	you are more responsive to visuals than text	**47**	you enjoy stories that blend science and humanity
23	you are logical and detached when evaluating stories	**48**	you find complex information mentally exhausting
24	you tend to daydream while reading narratives	**49**	you focus more on the moral of the story than the facts

25	you appreciate the historical significance of big events	**50**	you prefer light, entertaining content over serious narratives

Prompt: {

You are a participant in a psychology study about how people react to podcasts.

You just finished reading the following podcast transcript about the Apollo 11 Moon landing.

Imagine your genuine, human reaction. Respond as naturally as possible,

as if you were an ordinary person taking a survey after reading it.

Now answer the following question in JSON format and only include the number:

{{

"engagement": [number between 1 and 7]

}}

Question:

1. How engaging did you find this podcast episode?

(1 = Not at all engaging, 7 = Extremely engaging)

Your answer should sound like a real human rating — varied, spontaneous,

and not identical to others.}

Podcasts

Low flips. In July 1969, Apollo 11 was launched from Kennedy Space Center. The mission involved three astronauts: Neil Armstrong, Buzz Aldrin, and Michael Collins. The objective was to land on the Moon and return safely to Earth. At 9:32 a.m. local time, the Saturn V rocket lifted off as planned. The launch proceeded according to schedule, and the vehicle entered Earth orbit before continuing toward the Moon. Televised coverage documented each stage in real time. Viewers around the world observed the sequence of operations. The launch was monitored from multiple tracking stations positioned across several continents. Engineers analyzed telemetry to confirm stability in all major systems. Minor deviations in pressure and temperature were recorded but did not affect performance. Communication links were tested and verified throughout the ascent. After orbit insertion, the spacecraft performed a planned burn to begin translunar flight. The transition to deep-space operations was reported as nominal.

After three days of travel, the spacecraft entered lunar orbit. The crew completed required system checks and prepared for descent. A series of test transmissions confirmed continuous contact with Mission Control. Environmental readings indicated normal conditions inside the cabin. Armstrong and Aldrin transferred to the lunar module, named Eagle. Collins remained in the command module, Columbia, to maintain orbit and communications. Descent to the surface began on July 20. Several computer alarms appeared, but the guidance program continued to function. Armstrong took manual control and landed the module in a clear area known as the Sea of Tranquility. Contact with Mission Control was maintained throughout the procedure. Armstrong reported the landing with the message, “Houston, Tranquility Base here. The Eagle has landed.” Data logs recorded stable power output and consistent cabin pressure. Subsequent analysis confirmed that fuel margins were within an acceptable range at touchdown.

Several hours later, Armstrong exited the module and stepped onto the lunar surface. He described the event with the statement that became part of the record. Aldrin followed and conducted additional tasks as scheduled. Both astronauts collected rock samples, deployed instruments, and photographed the area. Surface operations followed the timeline established during simulation. Communication remained stable, and telemetry indicated no major issues. Collins continued orbiting the Moon while performing system checks and observations. He reported that all onboard systems were functioning normally. The astronauts documented material texture, soil depth, and shadow patterns as part of their field notes. Equipment positioning and instrument calibration were verified before departure.

The extravehicular activity lasted about two and a half hours. The astronauts then returned to the module, sealed the hatch, and prepared for ascent. Cabin systems were reset and waste containers secured. The launch from the lunar surface occurred without incident. Docking with the command module was completed successfully. Alignment was maintained within a few degrees of the target corridor. The crew then initiated the return trajectory to Earth. Mid-course corrections were conducted as scheduled and confirmed by tracking data. Re-entry occurred on July 24. The command module landed in the Pacific Ocean and was recovered by the U.S. Navy. The retrieval procedure followed standard safety protocol. The capsule was lifted onto the recovery ship and inspected for structural integrity.

Post-flight procedures included medical checks and a precautionary quarantine. All crew members were in good health. The collected samples were transferred to laboratories for analysis. Preliminary findings focused on mineral content, density, and isotopic variation. The mission met all primary objectives and confirmed the feasibility of manned lunar operations. Data from Apollo 11 provided reference material for later missions. The spacecraft and equipment were displayed in museums after debriefing and documentation. NASA released technical summaries to research institutions and partner agencies. Training programs for future missions incorporated modifications based on these findings.

Apollo 11 marked the first human landing on the Moon. It demonstrated coordinated engineering, accurate navigation, and controlled re-entry under defined conditions. The mission’s completion represented a technical milestone in space exploration. The event was observed by millions through global broadcast networks. Subsequent Apollo flights followed similar protocols based on these results. The historical record describes the operation as effective, efficient, and consistent with its stated goals. It remains a central example in studies of complex project management and human reliability. The event continues to serve as a reference point for discussions of human spaceflight, systems engineering, and international scientific cooperation.

High flips. The summer of 1969 unfolded under an atmosphere dense with both anticipation and unease, a season during which every announcement seemed to promise something larger than comprehension itself. Across newspapers and televisions, the repeated headlines proclaimed history not as a distant ideal but as an imminent event that demanded collective attention. At Kennedy Space Center, beneath a humid Florida sky, three men prepared to enter a machine that appeared at once magnificent and fragile, an object suspended between human invention and human limitation. Neil Armstrong, Buzz Aldrin, and Michael Collins strapped themselves into narrow seats, while the country, perhaps the entire world, held an almost involuntary breath. The Saturn V, standing immense and motionless, resembled less a vehicle than a monument to persistence and ambition, shaking the ground before any movement even began. When its engines finally ignited, the air trembled under a sound so immense that words, for a moment, seemed useless. The rocket lifted, unhurried yet unstoppable, leaving behind not only flame and vapor but a planet full of eyes fixed upward, caught between disbelief and pride.

Televisions flickered across cities and time zones as the vehicle disappeared into layered clouds, its image dissolving faster than the astonishment it left behind. Some viewers applauded instinctively; others watched in silence, uncertain whether what they were witnessing belonged to science or to myth. Children pointed toward the sky, their questions overlapping with the commentary that tried, inadequately, to explain what was happening. In Mission Control, engineers studied screens that glowed with data both reassuring and unreadable to the untrained eye. The numbers appeared steady, yet tension accumulated in the room like invisible pressure. When the phrase *orbit achieved* appeared, the applause that followed was brief, fragmented, and uncertain, as if no one yet trusted success.

Three long days extended across the black space between Earth and the Moon, each hour stretching into the next with almost measurable slowness. Outside, there was no sound, no color, only the steady progression of instruments recording the absence of anything familiar. Inside the capsule, conversation remained precise and minimal, its calm tone belying an awareness that risk existed in every motion. The Moon grew larger, an image once distant now occupying the windows like a reminder that imagination had always underestimated reality. Armstrong and Aldrin moved into the small lunar module, its walls close enough to touch on all sides, while Collins remained in orbit, solitary and silent. When the two modules separated, the smaller craft drifted away like an ember in darkness. The descent began unevenly, alarms interrupting the rhythm of procedure, each tone sharper than the last. Red lights flashed across the console; voices overlapped through static; heart rates spiked though none were mentioned aloud. Armstrong steadied the controls, ignoring the rising anxiety that filled the channel. "We're going to continue," he said, quietly enough to sound certain. Dust lifted beneath the landing legs; the horizon tilted; then came the words that reset history: "Houston, Tranquility Base here. The Eagle has landed."

The room in Houston erupted—not in chaos but in the structured release of tension that had held through too many hours. Men shouted, embraced, and covered their faces in disbelief that quickly resolved into relief. Around the planet, millions responded similarly, fear dissolving into exhilaration that bordered on disbelief. For one moment, the Cold War's division seemed to pause under the weight of a shared realization. Yet on the Moon, the procedures continued, the checklist advancing without sentiment. The landing had succeeded, but the work remained. Hours later, the hatch opened with a slowness that seemed deliberate. Armstrong descended the ladder, his movement cautious, almost hesitant, as though gravity itself required renegotiation. When his foot pressed into the dust, the transmission carried a pause so long it felt like silence before meaning. “That’s one small step for man, one giant leap for mankind.” The words, imperfect in form yet indelible in memory, circled the planet faster than comprehension could follow.

Aldrin followed, his description—“magnificent desolation”—merging contradiction into accuracy. The surface extended endlessly, both bright and dim, its stillness made heavier by the absence of sound. Together, they planted the flag, collected samples, and performed each movement with measured precision. Above them, Collins orbited alone, his isolation complete, his role unseen but essential. He later described the experience as simultaneously distant and profoundly connected, a paradox few could imagine. Below, his crewmates continued, part engineers, part explorers, each motion rehearsed yet newly uncertain.

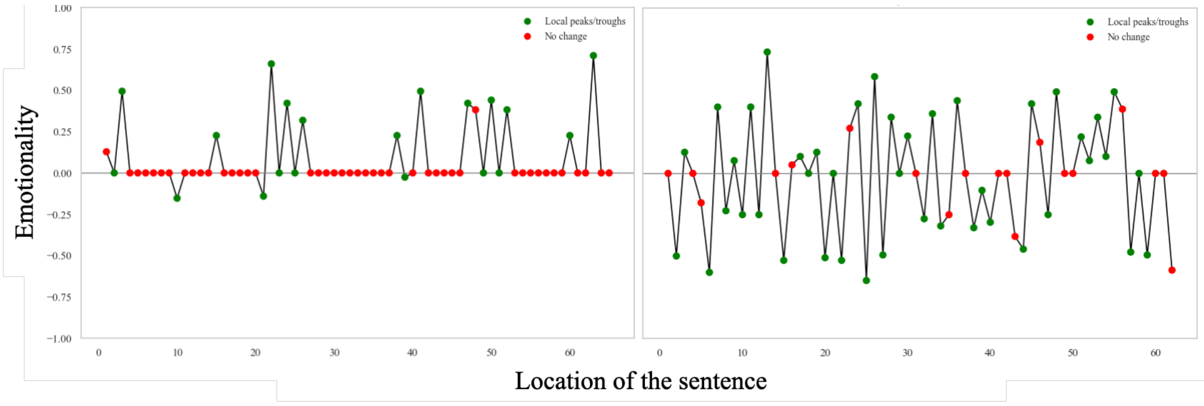
Back on Earth, entire cities vibrated with attention. Radios repeated transmissions until static replaced speech. Some listeners cried; others stood motionless, their faces turned toward a sky that revealed nothing. The mission, born of rivalry, had transformed into something unplanned and communal, a success that seemed to belong to everyone who had feared it might fail. The planet, for a brief time, appeared united by curiosity instead of caution.

Four days later, the command module re-entered the atmosphere in a burst of heat so intense that sensors saturated. The parachutes opened, white against the layered blue of the Pacific sky, and the capsule struck the water with a sound that was almost invisible in its simplicity. Recovery teams pulled the astronauts from the sea, their movements slow, methodical, and relieved. Flashbulbs exploded. Questions overlapped. The faces behind the glass of the quarantine unit reflected exhaustion more than triumph, though the distinction mattered little.

Across continents, celebrations varied but shared the same rhythm: spontaneous and uncertain. In one country, church bells rang; in another, radios played national anthems; in yet another, people simply looked upward. The phrase "If we can put a man on the Moon" appeared in headlines as though language itself had reached its upper limit. Children imitated countdowns on sidewalks, shouting numbers that dissolved into laughter. Scientists, back in their offices, began planning again almost immediately, their satisfaction already replaced by calculation.

Decades later, the recordings remain vivid in their imperfections—the static, the pauses, the heartbeats that microphones caught by accident. The story of Apollo 11 did more than record a journey; it redefined the boundaries of thought itself. It converted fear into coordination, rivalry into participation, and silence into the measured sound of human voices repeating numbers that meant survival. Even now, when the Moon appears above the horizon, pale and ordinary, it still carries that invisible residue of awe and discipline. The footprints, untouched by wind or time, remain as both evidence and question—what humans can do, and whether they should. Somewhere in that thin echo that travels between Earth and space, the voice persists—steady, imperfect, trembling slightly as it crosses history: "The Eagle has landed."

Figure I1. Emotionality flips across both conditions (Left: low emotionality flips, Right: high emotionality flips)



References

Adaval, R., Isbell, L. M., & Wyer, R. S. (2007). The impact of pictures on narrative- and list-based impression formation: A process interference model. *Journal of Experimental Social Psychology*, *43*(3), 352–364. https://doi.org/10.1016/j.jesp.2006.04.005

Adaval, R., & Wyer Jr., R. S. (1998). The Role of Narratives in Consumer Information Processing. *Journal of Consumer Psychology*, *7*(3), 207–245. https://doi.org/10.1207/s15327663jcp0703_01

Akpinar, E., & Berger, J. (2017). Valuable Virality. *Journal of Marketing Research*, *54*(2), 318–330. https://doi.org/10.1509/jmr.13.0350

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berger, J. (2011). Arousal Increases Social Transmission of Information. *Psychological Science*, *22*(7), 891–893. https://doi.org/10.1177/0956797611413294

Berger, J., Kim, Y. D., & Meyer, R. (2021). What Makes Content Engaging? How Emotional Dynamics Shape Success. *Journal of Consumer Research*, *48*(2), 235–250. https://doi.org/10.1093/jcr/ucab010

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

Berger, J., Moe, W. W., & Schweidel, D. A. (2023). What Holds Attention? Linguistic Drivers of Engagement. *Journal of Marketing*, *87*(5), 793–809. https://doi.org/10.1177/00222429231152880

Berman, R., Melumad, S., Humphrey, C., & Meyer, R. (2019). A Tale of Two Twitterspheres: Political Microblogging During and After the 2016 Primary and Presidential Debates. *Journal of Marketing Research*, *56*, 002224371986192. https://doi.org/10.1177/0022243719861923

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84.

Bojic, L., Kovacevic, P., & Cabarkapa, M. (2023). GPT-4 Surpassing Human Performance in Linguistic Pragmatics. https://arxiv.org/abs/2312.09545

Bublitz, M. G., Cross, S. N. N., Lteif, L., Nardini, G., Peracchio, L. A., Rank-Christman, T., & Woodrow, S. (2024). We Will Rise: How Stories Unite Social Movements. *Journal of the Association for Consumer Research*, *9*(1), 9–20. https://doi.org/10.1086/727837

Bulmer, M. G. (1979). *Principles of Statistics*. New York: Dover.

Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. (2009). *Reading Tea Leaves: How Humans Interpret Topic Models*. *Neural Information Processing Systems* (Vol. 32, p. 296).

Cuellar, J. E., Moreno-Martinez, O., Torres-Rodriguez, P. S., Pavlich-Mariscal, J. A., Mican-Castiblanco, A. F., & Torres-Hurtado, J. G. (2025). Trusting CHATGPT: how minor tweaks in the prompts lead to major differences in sentiment classification. https://arxiv.org/abs/2504.12180

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, *2*(3), 292–307. https://doi.org/10.1037/1082-989X.2.3.292

Escalas, J. E. (2004). Narrative processing: Building consumer connections to brands. *Journal of Consumer Psychology*, *14*(1–2), 168–180. https://doi.org/10.1207/s15327663jcp1401&2_19

Escalas, J. E. (2007). Self-Referencing and Persuasion: Narrative Transportation versus Analytical Elaboration. *Journal of Consumer Research*, *33*(4), 421–429. https://doi.org/10.1086/510216

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511790942

Green, M., & Brock, T. (2000). The Role of Transportation in the Persuasiveness of Public Narrative. *Journal of personality and social psychology*, *79*, 701–21. https://doi.org/10.1037/0022-3514.79.5.701

Green, M. C. (2004). Transportation Into Narrative Worlds: The Role of Prior Knowledge and Perceived Realism. *Discourse Processes*, *38*(2), 247–266. https://doi.org/10.1207/s15326950dp3802_5

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., et al. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794

Heimbach, I., & Hinz, O. (2016). The impact of content sentiment and emotionality on content virality. *International Journal of Research in Marketing*, *33*(3), 695–701. https://doi.org/10.1016/j.ijresmar.2016.02.004

Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., & Schoegel, M. (2019). Detecting, Preventing, and Mitigating Online Firestorms in Brand Communities. *Journal of Marketing*, *83*(3), 1–21. https://doi.org/10.1177/0022242918822300

Hwang, E. H., & Lee, S. (2024). A Nudge to Credible Information as a Countermeasure to Misinformation: Evidence from Twitter. *Information Systems Research*. https://doi.org/10.1287/isre.2021.0491

Krause-Galoni, R., & Rucker, D. D. (2024). Fact versus Fiction: Blending Storytelling with Product Attributes in Advertisements. *Journal of the Association for Consumer Research*, *9*(1), 58–70. https://doi.org/10.1086/727835

Lee, J. K. (2021). Emotional Expressions and Brand Status. *Journal of Marketing Research*, *58*(6), 1178–1196. https://doi.org/10.1177/00222437211037340

Li, Y., & Xie, Y. (2020). Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement. *Journal of Marketing Research*, *57*(1), 1–19. https://doi.org/10.1177/0022243719881113

Lin, Y., Yao, D., & Chen, X. (2021). Happiness Begets Money: Emotion and Engagement in Live Streaming. *Journal of Marketing Research*, *58*(3), 417–438. https://doi.org/10.1177/00222437211002477

Martin, N. S., VanBergen, N., Poehlman, T. A., & Adaval, R. (2024). Connecting the Past and the Present: How Narrative-Based Representations Affect Perceptions of Brand Heritage. *Journal of the Association for Consumer Research*, *9*(1), 71–82. https://doi.org/10.1086/727834

Melumad, S., Inman, J. J., & Pham, M. T. (2019). Selectively Emotional: How Smartphone Use Changes User-Generated Content. *Journal of Marketing Research*, *56*(2), 259–275. https://doi.org/10.1177/0022243718815429

Mohammad, S., & Bravo-Marquez, F. (2017). *Emotion Intensities in Tweets* (p. 77). https://doi.org/10.18653/v1/S17-1007

Mosley, B., Schweidel, D. A., & Zhang, K. (2024). When Connection Turns to Anger: How Consumer–Brand Relationship and Crisis Type Moderate Language on Social Media. *Journal of Consumer Research*, *50*(5), 907–922. https://doi.org/10.1093/jcr/ucad027

Mousavi, R., & Gu, B. (2023). Resilience Messaging: The Effect of Governors' Social Media Communications on Community Compliance During a Public Health Crisis. *Information Systems Research*. https://doi.org/10.1287/isre.2021.0599

Packard, G., Berger, J., & Boghrati, R. (2023). How Verb Tense Shapes Persuasion. *Journal of Consumer Research*, *50*(3), 645–660. https://doi.org/10.1093/jcr/ucad006

Phillips, B. J., & McQuarrie, E. F. (2010). Narrative and Persuasion in Fashion Advertising. *Journal of Consumer Research*, *37*(3), 368–392. https://doi.org/10.1086/653087

Ransbotham, S., Lurie, N. H., & Liu, H. (2019). Creation and Consumption of Mobile Word of Mouth: How Are Mobile Reviews Different? *Marketing Science*, *38*(5), 773–792. https://doi.org/10.1287/mksc.2018.1115

Ravichandran, T., & Deng, C. (2023). Effects of Managerial Response to Negative Reviews on Future Review Valence and Complaints. *Information Systems Research*, *34*(1), 319–341. https://doi.org/10.1287/isre.2022.1122

Rivera, M., Jiang, C., & Kumar, S. (2023). Seek and Ye Shall Find: An Empirical Examination of the Effects of Seeking Real-Time Feedback on Employee Performance Evaluations. *Information Systems Research*. https://doi.org/10.1287/isre.2021.0130

Tellis, G. J., MacInnis, D. J., Tirunillai, S., & Zhang, Y. (2019). What Drives Virality (Sharing) of Online Digital Content? The Critical Role of Information, Emotion, and Brand Prominence. *Journal of Marketing*, *83*(4), 1–20. https://doi.org/10.1177/0022242919841034

Toubia, O., Berger, J., & Eliashberg, J. (2021). How quantifying the shape of stories predicts their success. *Proceedings of the National Academy of Sciences*, *118*(26), e2011695118. https://doi.org/10.1073/pnas.2011695118

Valenzuela, A., & Galli, M. (2024). Intimate Transportation: The Persuasive Role of Personal Narratives in Online Reviews. *Journal of the Association for Consumer Research*, *9*(1), 83–94. https://doi.org/10.1086/727832

Vamvourellis, D., & Mehta, D. (2025). Reasoning or Overthinking: Evaluating Large Language Models on Financial Sentiment Analysis. https://arxiv.org/abs/2506.04574

van Laer, T., de Ruyter, K., Visconti, L. M., & Wetzels, M. (2014). The Extended Transportation-Imagery Model: A Meta-Analysis of the Antecedents and Consequences of Consumers' Narrative Transportation. *Journal of Consumer Research*, *40*(5), 797–817. https://doi.org/10.1086/673383

van Laer, T., Edson Escalas, J., Ludwig, S., & van den Hende, E. A. (2019). What Happens in Vegas Stays on TripAdvisor? A Theory and Technique to Understand Narrativity in Consumer Reviews. *Journal of Consumer Research*, *46*(2), 267–285. https://doi.org/10.1093/jcr/ucy067

Wang, Y., Xu, X., Kuchmaner, C. A., & Xu, R. (2023). But it was supposed to be healthy! How expected and actual nutritional value affect the content and linguistic characteristics of online reviews for food products. *Journal of Consumer Psychology*, *33*(4), 743–761. https://doi.org/10.1002/jcpy.1376

Yazdani, E., Gopinath, S., & Carson, S. (2018). Preaching to the Choir: The Chasm Between Top-Ranked Reviewers, Mainstream Customers, and Product Sales. *Marketing Science*, *37*(5), 838–851. https://doi.org/10.1287/mksc.2018.1101