Getting the Recipe Right: How Content Combinations Drive Social Media Engagement Behaviors

Lapresta-Romero, S., Becker, L., Hernández-Ortega, B., Terho, H., & Franco, J. L. (2024). Getting the recipe right: How content combinations drive social media engagement. *Journal of Interactive Marketing*, doi/10.1177/10949968241251694.

Abstract

Social media has become a key touchpoint in contemporary customer journeys. Consequently, prior studies have investigated how social media content drives outcomes. However, much of this research has focused on the design of individual, isolated content elements, paying limited attention to how individuals respond to their holistic combinations. Drawing on multimodality, this study investigates how combinations of content elements drive social media engagement behaviors (SMEBs), a critical social media outcome. Through a fuzzy-set qualitative comparative analysis with 516 Instagram stories, the findings reveal four content element configurations that can drive high SMEBs: the loud, the informative, the affective, and the relational. These findings contribute to the literature by demonstrating that multiple configurations of content elements can simultaneously drive SMEBs, thus challenging the dominant view in the literature, which has focused on the effectiveness of isolated elements on diverse outcomes.

Keywords: social media, social media engagement behaviors, engagement, content elements, social media content, fsQCA, multimodality

The multimodal communication capabilities of social media are radically changing how firms can engage with customers (Dhaoui and Webster 2021). As increasingly central touchpoints for contemporary customer journeys, social media platforms have become critical tools for digital marketers (Kumar et al. 2016), such that more than 90% of marketers anticipated positive returns on their social media marketing investments in 2023 (HubSpot 2022). Among the customer-related social media outcomes they seek, one of the most critical is social media engagement behaviors (SMEBs) (Liadeli, Sotgiu, and Verlegh 2022), defined as a "customer's behavioral manifestations that have a social media focus, beyond purchase, resulting from motivational drivers" (Dolan et al. 2016, p. 265). Given that SMEBs shape customer equity, loyalty, and firm growth and performance (e.g., Barger, Peltier, and Schultz 2016; Cao et al. 2021; van Doorn et al. 2010), both academics and marketers are striving to better understand their drivers (Pezzuti, Leonhardt, and Warren 2021; Rietveld et al. 2020; Villarroel Ordenes et al. 2019).

A key driver of SMEBs is social media content. Social media research shows that content elements such as emotional expressiveness, visual patterns, and words of certainty influence SMEBs (e.g., Farace et al. 2020; Holiday et al. 2023; Pezzuti, Leonhardt, and Warren 2021; Villarroel Ordenes et al. 2019). Among these studies, though, we note two major limitations. First, despite the multimodal nature of social media (Villarroel Ordenes et al. 2019) and the efficacy of using multiple content elements in conveying firm messages (Holiday et al. 2023), research on multimodal digital communication is scarce (Grewal et al. 2022). Social media content often incorporates music, images, and various symbols, engaging customers' sight, hearing, and emotionality. However, social media research tends to investigate the unique effects of specific content elements on SMEBs, such as visual or linguistic (Mulier, Slabbinck, and Vermeir 2021; Pezzuti, Leonhardt, and Warren 2021) (see Table 1), while failing to integrate other types of elements.

Second, social media research pays scant attention to the combinatory effects of multiple content elements on SMEBs. This is a significant limitation because content elements jointly influence diverse outcomes, rather than acting independently (Grewal et al. 2022). In other words, customers do not respond to elements in isolation but rather to their combinations (Bleier, Harmeling, and Palmatier 2019; Grewal et al. 2022; Holiday et al. 2023). Accordingly, studies of social media need to shift their focus from examining customers' responses to individual elements toward understanding how holistic combinations of elements drive SMEBs.

This study aims *to examine how combinations of multiple content elements drive SMEBs*. We draw on multimodality (e.g., Kress 2010) and various marketing literature streams to identify six key content elements applicable to social media: visual, auditory, linguistic, symbolic, social, and emotional. We contend that the fit of these elements determines their effectiveness in generating SMEBs. Using a fuzzy-set qualitative comparative analysis (fsQCA), we examine a data set of 516 Instagram stories posted by a tourism resort operator—an experiential context that promotes consumer-brand interactions on social media with the aim of fostering strong relationships and engagement behaviors (Mariani, Di Felice, and Mura 2016; Yousaf et al. 2021). Instagram stories are characterized by their fleeting nature and multiple, rich content elements (Chen and Cheung 2019); as such, the data set is well-suited as a representation of social media content, aligning seamlessly with the chosen multimodal research focus. The findings reveal four effective content configurations that enhance SMEBs: the loud, the informative, the affective, and the relational. These results show that configurations of content elements, rather than isolated elements, drive SMEBs.

We make two novel contributions to social media research. First, unlike most research to date focusing on a limited number of content elements, we identify key content elements in social media (i.e., visual, auditory, linguistic, symbolic, social, and emotional). By doing so, we offer a more comprehensive view of the content elements applicable to any social media as a multimodal means of communication. Moreover, we advance social media research by demonstrating that content elements are best approached and studied holistically. Rather than focusing solely on the linear effects of individual content elements on SMEBs, we demonstrate that holistic combinations of content elements provide powerful means to predict SMEBs, with the efficacy of each element dependent on its association with the others. Second, we identify four alternative configurations of content elements that explain high SMEBs. These configurations highlight that there is no single way to design effective content to drive SMEBs; instead, several configurations can simultaneously drive strong SMEBs.

Literature Review

Extant View on How Content Elements Drive SMEBs

Defined as customer actions motivated by factors beyond the purchase and in relation to a firm's social media presence (Dolan et al. 2016; van Doorn et al. 2010), SMEBs are typically operationalized by metrics such as the number of likes, comments, or shares in response to social media content (Deng et al. 2021; McShane, Pancer, and Poole 2019; Pezzuti, Leonhardt, and Warren 2021; Zhao et al. 2023). These typical indicators of social media success (Liadeli, Sotgiu, and Verlegh 2022) can yield positive outcomes, such as brand awareness (Swani et al. 2017), customer equity (van Doorn et al. 2010), loyalty (Cao et al. 2021), purchases, firm growth, and performance (Barger, Peltier, and Schultz 2016; Lin, Swarna, and Bruning 2017; Santini et al. 2020). Accordingly, both academics and marketers increasingly seek to understand how social media content should be designed to drive SMEBs.

Table 1 presents an overview of key studies investigating the relationship between content elements and SMEBs. While these studies have provided a cumulative body of research on the nature of content elements that drive engagement, our review identifies two central limitations of this extant research. First, existing research has focused heavily on certain content elements within social media, primarily visual and linguistic (e.g., Deng et al. 2021; Dhanesh, Duthler, and Li 2022). However, given the evolution of social media into a multimodal means of communication (Grewal et al. 2022; Villarroel Ordenes et al. 2019), there is a need to incorporate a more diverse set of content elements. Examining multiple elements allows us to provide more comprehensive explanations regarding the effectiveness of social media content.

Second, most studies investigate the net effects of individual content elements on SMEBs, applying correlation-based methods such as regression analysis to estimate direct and interaction effects between visual and linguistic elements (Farace et al. 2020; Holiday et al. 2023; Villarroel Ordenes et al. 2019). These studies treat relationships between variables as symmetric, leading to contradictory findings for similar elements across different studies (Ceylan, Diehl, and Wood 2023; Deng et al. 2021; Li and Xie 2020). Only a few exceptions seek to deepen understanding of the overall impact of how combinations of content elements drive SMEBs, employing techniques such as decision trees. For example, Yu and Egger (2021) use support vector machine and random forest methods to investigate the influence of color composition on engagement rates for different types of Instagram posts. Surucu-Balci, Balci, and Yuen (2020) apply decision trees to identify how linguistic elements, content type (e.g., advertisement, celebration, company news), and vividness work together to induce SMEBs. Both of these studies seek to identify the most effective single combination of content elements. However, we challenge this notion of a singular pathway to SMEBs and suggest that multiple simultaneous combinations of elements can drive SMEBs.

Table 1: Key Studies on the Relationship between Content Elements and SMEBs.

Study	Content type	Data collection (tool)	Method of analysis	Studied effect	Studied outcome	Key findings
Aleti et al. (2019)	Text-based tweets	LIWC	Latent class analysis	Linear effects of individual elements	Word of mouth (retweets)	The linguistic style of tweets drives word of mouth on Twitter. Specifically, narrative styles generate higher word of mouth, whereas emotional styles are not effective.
Deng et al. (2021)	Text-based tweets	LIWC	Multiple regression	Linear effects of individual elements	Brand engagement (likes and retweets)	Linguistic features influence brand engagement: post length, language complexity, visual complexity, emotional elements, interpersonal elements, and multimodal elements. Linguistic features that enable (prevent) central or peripheral route processing have a positive (negative) impact on brand engagement.
Dhanesh, Duthler, and Li (2022)	Image-based posts	Human coders	t-test	Linear effects of individual elements	Social media engagement (likes and comments)	Narrative, interactive features of distance and point of view features of images drive engagement on Facebook and Instagram; compositional characteristics of framing of images increase engagement on Instagram.
Farace et al. (2020)	Text- and image-based tweets	Human coders and LIWC	Poisson regression model	Linear and interaction effects of individual elements	Likes and retweets	The combination of regular visual patterns with a caption that conveys motion positively influences likes and retweets.
Hernández-Ortega et al. (2020)	Image-based posts	LIWC and human coders	Generalized method of moments regression	Linear effects of individual elements	Popularity (organic reach)	Affective image attributes positively influence the popularity of content on Facebook; cognitive image attributes do not influence popularity.
Holiday et al. (2023)	Text- and image-based posts	Facial expression analysis, LIWC, human coders	Negative binomial regression	Linear and interaction effects of individual elements	Engagement (likes, comments, views)	The combination of facial and textual emotional expressiveness positively influences comments, but not likes or views.
Li and Xie (2020)	Text- and image-based posts and tweets	LIWC and Google Cloud Vision	Bivariate zero-inflated negative binomial regression	Linear effects of individual elements	Social media engagement (likes and retweets)	Image content drives engagement through product categories. High-quality, professionally shot images lead to better engagement. The impact of color differs by product category. The appearance of a human face and image–text fit increase engagement on Twitter but not on Instagram.
Luangrath, Peck, and Barger (2017)	Text-based tweets, and posts	Text analysis markup system analyzer, human coders	Exploratory: frequencies and percentages	Proposed effects of individual elements	Brand–consumer relationship (engagement)	A framework outlines the antecedents and consequences of using textual paralanguage (auditory textual, tactile textual, and visual textual). The brand–consumer relationship is a consequence of textual paralanguage.

Mulier, Slabbinck, and Vermeir (2021)	Video-based posts	Facebook A/B split test and survey	Z-test and Tobit regression	Linear effects of individual elements	Engagement (likes, clicks, comments, shares)	Mobile vertical (vs. horizontal) video ads enhance consumer interest and engagement.
Overgoor et al. (2022)	Image-based posts	Data mining methods	Negative binomial regression	Linear effects of individual elements	Likes	Feature complexity and design complexity of firm-generated imagery influence consumer liking.
Pezzuti, Leonhardt, and Warren (2021)	Text-based posts	LIWC	Negative binomial regression	Linear effects of individual elements	Consumer engagement (likes, shares, comments)	Words of certainty in brand messages on Facebook lead to higher levels of engagement through perceptions of power.
Rietveld et al. (2020)	Image-based posts	Machine learning	Negative binomial regression	Linear effects of individual elements	Engagement (likes, comments)	Visual-emotional and informative appeals embedded in brand-generated content increase customer engagement.
Surucu-Balci, Balci, and Yuen (2020)	Text- and image-based tweets	Human coders	Decision tree	Effective combination of elements	Stakeholder engagement (likes, retweets, comments)	A combination of fluency of tweets, the tangibility of company resources in the tweet, the vividness level, the content type, the presence of a link, and the presence of a call-to-action drive stakeholder engagement rate.
Villarroel Ordenes et al. (2019)	Text- and image-based tweets, and posts	Human coders and machine learning	Negative binomial and Poisson regression	Linear and interaction effects of individual elements	Shares	The use of rhetorical styles and cross-message compositions boosts consumer message sharing, and the presence of visuals increases the capacity to account for message sharing.
Yu, Xie, and Wen (2020)	Image-based posts	Data mining and pictorial content analysis	Multiple regression	Linear effects of individual elements	Popularity (likes, comments)	Brighter and more saturated destination images on Instagram influence popularity.
Yu and Egger (2021)	Image-based posts	Google Cloud Vision	Support vector machine and random forest	Effective combination of elements	Engagement (likes, comments)	In driving engagement, effective color compositions related to image hues vary depending on the type of images (e.g., mountains and water or urban views).
Zhao et al. (2023)	Image-based posts	Object detection techniques	Multivariate regression	Linear effects of individual elements	Engagement (likes, shares)	Image richness positively influences emotional and behavioral engagement but negatively impacts cognitive engagement.
This study	Image- and video-based stories	Machine learning, human coders	fsQCA	Multiple effective combinations of elements	SMEBs (comments, shares)	Examining how combinations of content elements (visual, auditory, linguistic, symbolic, social, and emotional) drive SMEBs. Multiple effective combinations of these content elements, rather than elements in isolation, drive SMEBs.

Note: LIWC = Linguistic Inquiry and Word Count dictionary.

Multimodal Content in Social Media

This study builds on multimodality, a theoretical approach within the field of communication that explores the various ways in which meaning is created and conveyed (Kress 2010). This approach acknowledges that communication transcends traditional linguistic forms as individuals employ a variety of semiotic resources, including images, colors, shapes, speech, gestures, and writing, among others, to build more comprehensive meanings (Kress 2010; Poulsen and Kvåle 2018). These semiotic resources offer unique ways of transmitting significance, thereby enhancing the overall effectiveness of the communicative process (Kress 2010; 2011).

One of the key arguments of multimodality emphasizes how the dynamic interplay among diverse semiotic resources constructs meanings in complex and nuanced ways (Grewal et al. 2022; Kress 2010). In essence, these resources are not independent; instead, they interact and converge within a communicative context, giving rise to ensembles that significantly influence the construction and expression of meaning (Jewitt 2013). Thus, semiotic resources consistently appear in combinations, deliberately crafted with a focus on their interrelations and interactions (Bezemer and Kress 2016).

Social media is a multimodal means of communication that often incorporates diverse content elements such as images, text, emojis, GIFs, and numbers (Grewal et al. 2022; Poulsen and Kvåle 2018; Villarroel Ordenes et al. 2019). This variety enables social media marketers to employ numerous semiotic resources in their interactions with customers, enabling them to effectively communicate and convey nuanced meaning (Kress 2010; Poulsen and Kvåle 2018). In this context, recent studies suggest that the simultaneous combination of multiple content elements, rather than a singular element, determines the impact of communication on important outcomes such as SMEBs (Ceylan, Diehl, and Wood 2023; Grewal et al. 2022).

The present research employs multimodality (Grewal et al. 2022; Kress 2010) as a comprehensive theoretical approach. This requires breaking down the object of study (i.e., social media content) into its component parts (i.e., content elements) to understand how these parts work together to achieve a particular outcome (i.e., SMEBs) (Flewitt, Price, and Korkiakangas 2019). Consequently, we view content elements in social media as semiotic resources that, when used in combination, have the potential to convey meaning in communication and lead to desired outcomes. We discuss these content elements next.

Content Elements in Social Media

Following the multimodality approach, we expand our review beyond social media research, which is often limited to visual and linguistic elements (see Table 1), and endeavor to identify content elements that can be employed in social media. Our exploration encompasses broader marketing literature streams, delving into retail and service atmospherics (e.g., Baker et al. 2002), customer experience (e.g., Bleier, Harmeling, and Palmatier 2019), sensory marketing (e.g., Krishna 2012), and online environment design (e.g., Bashirzadeh, Mai, and Faure 2021). Drawing on these literature streams, we identify content elements that represent key semiotic resources capable of enriching communication in any social media (e.g., the servicescape literature acknowledges the influence of social elements on customer outcomes; Baker et al. 2002). Specifically, we identify six types of content elements that collectively address stylistic, technical, and meaning-making resources: visual, auditory, linguistic, symbolic, social, and emotional elements.

Visual elements

Visual elements encompass resources that are employed to immediately appeal to the sense of sight, typically requiring less cognitive effort to process compared to words (McShane, Pancer, and Poole 2019). Previous research cites several types of visual elements:

colors (Brakus, Schmitt, and Zarantonello 2009; Lam et al. 2011) and specific color characteristics such as hue, saturation, and brightness (HSB) (Chi, Pan, and Huang 2021; Hsieh et al. 2018); lighting (Roggeveen, Grewal, and Schweiger 2020); visual variation and complexity (Li, Shi, and Wang 2019; Tuch et al. 2009); and architecture, including shapes, design, and layout (Baker et al. 2002; Demoulin and Willems 2019; Roggeveen, Grewal, and Schweiger 2020).

Visual resources related to colors (e.g., HSB) and visual richness (e.g., diversity in shapes, objects, and patterns) are of particular relevance to social media. Prior research suggests that the effects of colors and visual richness on SMEBs depend on the context; for instance, the impacts of HSB appear mixed (Li and Xie 2020; Yu, Xie, and Wen 2020; Yu and Egger 2021), while visual richness exhibits nonlinear effects on SMEBs (Overgoor et al. 2022).

Auditory elements

Auditory elements are resources perceived through the sense of hearing, designed to convey meaning independently of semantic content and often evoked by sounds, music, or voices (Oakes and North 2008). Previous studies identify auditory elements such as noise (Bitner 1992; Demoulin and Willems 2019), music (Baker et al. 2002), context-specific sounds (e.g., coin or machine sounds in casinos) (Lam et al. 2011), and ambient sounds, jingles, and voice (Krishna 2012), as well as attributes of certain auditory elements, such as speech, volume, pitch, tempo, tonality, and texture (Barcelos, Dantas, and Sénécal 2018; Krishna 2012).

Music and speech are prevalent auditory resources in social media. The presence and speed of music (e.g., tempo) and voice (e.g., speech rate) can influence customer message processing, information load, and the time needed to process the message (Hahn and Hwang 1999; Rodero 2016). Studies on speech rates in online environments show mixed results

(Yang, Yang, and Zhou 2022). According to Rodero (2016), any stimulus must be sufficiently fast or dynamic to capture listeners' attention, but it should also maintain a moderate speed to avoid hindering proper processing.

Linguistic elements

Linguistic elements represent text-based resources (Kim et al. 2021; Kim and Lennon 2008). Previous research has examined various aspects of linguistic elements, including text length (Chevalier and Mayzlin 2006), topics of conversation (Swaminathan et al. 2022), information quality (Filieri et al. 2021), linguistic style (Liebrecht, Tsaousi, and van Hooijdonk 2021), inter-letter spacing, font size, and typography (Grobelny and Michalski 2015).

In the context of social media, alongside text, the inclusion of non-alphanumeric characters (e.g., #, @) involves the use of hashtags and mentions. These elements are key linguistic resources that can impact both clarity and the cognitive effort required for information processing (Deng et al. 2021; McShane, Pancer, and Poole 2019). Previous research presents mixed findings regarding the influence of these elements on SMEBs, again suggesting context-dependent effects (Deng et al. 2021; Li and Xie 2020; Surucu-Balci, Balci, and Yuen 2020).

Symbolic elements

Symbolic elements involve the use of symbols and signs as resources to communicate abstract concepts to customers (Chen, Huarng, and González 2022). The elements carry connotations beyond their literal meaning, thus adding significance and facilitating essential contextual communication (Luangrath, Peck, and Barger 2017). Prior literature has identified examples of symbolic elements such as logos and signage (Bitner 1992; Brakus, Schmitt, and Zarantonello 2009), artifacts and style (Bitner 1992; Lam et al. 2011), and, specific to online environments, animated images (Laroche et al. 2022), avatars (Azer et al. 2023), emoticons, emojis, and GIFs (Bashirzadeh, Mai, and Faure 2021; Boutet et al. 2021).

Emojis and GIFs are particularly relevant resources in social media content that can lead to higher SMEBs (Deng et al. 2021; Surucu-Balci, Balci, and Yuen 2020). Interactions among symbolic elements (i.e., employing GIFs and emojis jointly or separately) can influence message outcomes such as click-through and (un)subscription rates (Bashirzadeh, Mai, and Faure 2021).

Social elements

Social elements refer to resources related to people in the environment (Ponsignon, Durrieu, and Bouzdine-Chameeva 2017), such as employees' behaviors, performance, personality, and physical appearance (Baker et al. 2002; Lam et al. 2011); the presence of other individuals and customers (Demoulin and Willems 2019; Li, Shi, and Wang 2019; Ponsignon, Durrieu, and Bouzdine-Chameeva 2017; Roggeveen, Grewal, and Schweiger 2020); and interactions between employees and customers (Baker et al. 2002).

In social media, resources related to people reflect social presence and social interactions (Wang 2020). Social presence (e.g., whether people appear in content) and social interactions (e.g., direct gestures or gaze at customers) may influence perceptions and drive customer behaviors (Dhanesh, Duthler, and Li 2022). Intuitively, we might expect that the presence of people and direct gaze in social media content would result in greater SMEBs; however, research has provided no conclusive findings (Dhanesh, Duthler, and Li 2022; Li and Xie 2020).

Emotional elements

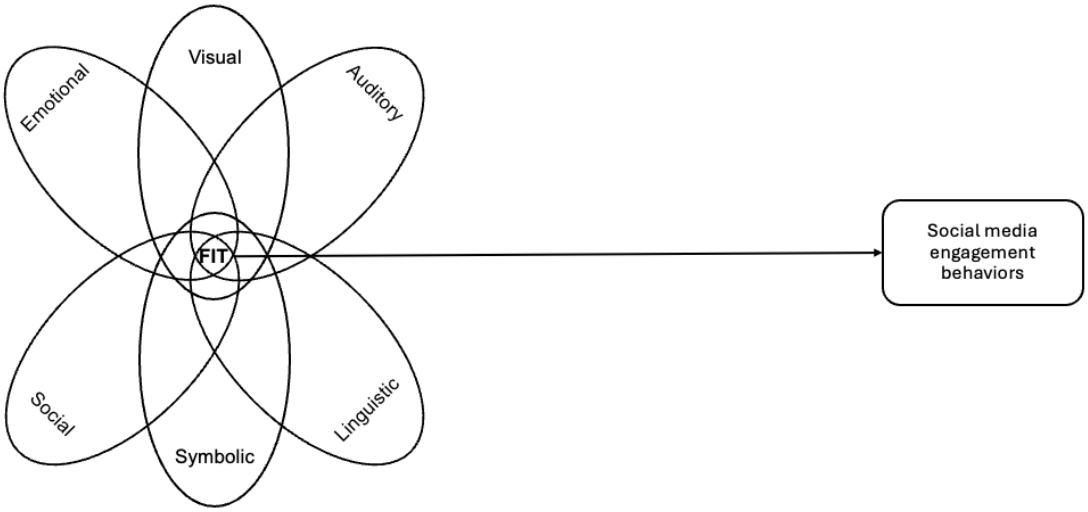
Emotional elements encompass resources that convey emotions, affect, and feelings, aiming to elicit emotional reactions (Rietveld et al. 2020). Some research has explored the effects of social media content that evokes emotions, feelings, or moods (Sykora et al. 2022). Most studies indicate that emotional elements result in positive outcomes (Alamäki, Pesonen, and Dirin 2019). For example, direct expressions of emotions such as "blew my mind" convey a strong sense of surprise (Sykora et al. 2022). Other emotional resources include the representation of emotions conveyed through words, hashtags, and emojis (Boutet et al. 2021; Klostermann et al. 2018; Ludwig et al. 2013).

In social media, any message that conveys feelings, moods, and emotions holds significance as a valuable resource. For example, firms can employ emotional resources in images (e.g., a picture of a heart) to encourage customers to express their feelings by engaging with the message (Swani, Milne, and Brown 2013) as a form of SMEB (Deng et al. 2021; Rietveld et al. 2020; Swani et al. 2017). Nonetheless, important differences remain in the findings obtained across different platforms and contexts (Aleti et al. 2019; Holiday et al. 2023; Swani, Milne, and Brown 2013).

Conceptual Framework: How Combinations of Content Elements Drive SMEBs

By integrating the key points from prior sections, we propose a conceptual framework that highlights the role of content element combinations in driving SMEBs (see Figure 1). These elements represent semiotic resources used in combination to elicit specific outcomes. They may vary in intensity, and their different combinations can collectively form distinct content configurations. Each configuration is characterized by the presence of tightly interconnected content elements (e.g., auditory, visual) and the absence of others (e.g., linguistic, symbolic). In social media, no individual element can be understood in isolation from others; instead, multiple combinations of content elements can drive the target outcome. Thus, social media content should be designed holistically, considering the fit among the six types of content elements to drive SMEBs.

Figure 1: Conceptual Framework.



Method

Fuzzy-Set Qualitative Comparative Analysis (fsQCA)

To determine how different content elements interact to drive SMEBs, we use *fuzzy-set qualitative comparative analysis (fsQCA)*, a set-theoretic approach well-suited to analyzing complex causal relationships among multiple conditions, as demanded by multimodality (Ragin 2008). It distinguishes between conditions that are connected or unconnected to an outcome of interest by testing multiple combinations, reflecting the notion that a given outcome can result from different combinations of sets of conditions (Fiss 2011; Ragin 2000).

FsQCA operates on the assumptions of conjunctural causation, equifinality, and causal asymmetry, and it distinguishes between necessary and sufficient conditions for an outcome to occur (Schneider and Eggert 2014). *Conjunctural causation* assumes that the interplay of conditions, rather than isolated conditions, drives the outcome of interest (Schneider and Eggert 2014). *Equifinality* assumes that multiple combinations of conditions can be equally effective in driving an outcome of interest (Meyer, Tsui, and Hinings 1993;

Pappas and Woodside 2021). *Causal asymmetry* assumes that while the presence of certain conditions may lead to an outcome, their absence does not necessarily result in the absence of the outcome (Fiss 2011). Finally, the conditions driving an outcome can be either *necessary* or *sufficient*. Necessary conditions must be present for the outcome to occur; sufficient conditions imply that the outcome arises whenever those conditions are present (Frösén et al. 2016; Salonen et al. 2021).

FsQCA offers several advantages over traditional models such as regression analysis, which primarily focuses on estimating the net effects of independent variables on a dependent variable (Mahoney and Goertz 2006). While regression analysis can be useful in various contexts, it is ill-suited to studying causally complex relationships for two main reasons. First, it offers limited insights into conjunctural causation by examining only a few moderating effects at a time. Second, it cannot address equifinality, handle causal asymmetry, or distinguish between necessary and sufficient conditions (Pappas and Woodside 2021). In contrast, our multimodality-focused research requires the use of more advanced methods such as fsQCA or tree-based models designed to analyze such causal complexity in explaining an outcome. Lindner, Puck, and Verbeke (2022, p. 1312) note that these methods share a similar logic but differ in emphasis, such that "QCA focuses on parsimony and on maximizing the predictive power of a branch of a tree (Fainshmidt et al. 2020), while tree-based models emphasize the simplicity of an explanation (Bauer and Kohavi 1999)". Consequently, we build on the qualitative-focused fsQCA, enabling us to develop detailed yet parsimonious explanations of how social media content elements, acting as causal conditions, jointly explain SMEBs as the outcome of interest (Fiss 2011; Ragin 2000).

We conduct the data analysis in line with the established fsQCA guidelines, which comprise six steps: (1) calibration and transformation of conditions, (2) construction of the truth table, (3) identification of necessary and (4) sufficient conditions, (5) interpretation of results, and (6) assessment of predictive validity (Frösén et al. 2016; Pappas and Woodside 2021; Salonen et al. 2021). For clarity, here in the Method section we describe the data collection process and operationalization of conditions, the calibration and transformation of conditions, and the construction of the truth table. Then, in the Results section, we report the identification of necessary and sufficient conditions, interpretation of results, and robustness checks.

Data Collection and Operationalization of Conditions

We obtained access to a relevant data set from a snow tourism resort operator in a southern European country. Our focus on the tourism sector stems from previous research demonstrating that social media within this sector has evolved into a crucial means of communication for fostering consumer engagement behaviors, such as information search, experience sharing, and value co-creation with brands (Mariani et al. 2016; Yousaf et al. 2021). The data set comprises all firm-generated Instagram stories published between February and March 2019 by one of the operator's resorts, totaling 516 sets of photos and videos. Instagram stories can display various combinations of different content elements (Chen and Cheung 2019), making it an interesting setting for analyzing multimodality in social media.

The unit of analysis is each story (n = 516), but we also extract all frames from video-based stories, with one second representing one frame, so that the data set contains 2,333 frames.[1] By identifying all elements present in each frame, we capture all information that customers can perceive during a video. This approach enables us to examine the presence of a content element if it appears, even if only in one frame.

[1] In Instagram stories, videos are limited to a maximum duration of 15 seconds, equivalent to 15 frames. The minimum duration is 3 seconds, or 3 frames.

Table 2 presents the operationalization of the conditions (i.e., content elements) and the outcome of interest (i.e., SMEBs). Similar to prior social media studies (Dolan et al. 2016), we operationalize SMEBs as the shares and comments received by each Instagram story. These contributions represent strong forms of engagement behaviors suitable for analyzing the outcome of any social media content (Liadeli, Sotgiu, and Verlegh 2022; Robiady, Windasari, and Nita 2021).[2] This operationalization involves three steps. First, we calculate share and comment rates by dividing the number of shares and comments of the story, respectively, by the total number of impressions it has received (i.e., the number of times that the story appeared on a customer's screen). These rates mitigate biases from other variables that cannot be controlled (e.g., time of publication) and are similar to the click-through rate indicator used in advertising.

Second, we define the thresholds for share and comment rates by considering historical SMEB data (i.e., shares and comments) obtained from stories published by other snow resorts in the same country over the past year. The firm employs this data to compare and assess customer reactions to its social media publications, ultimately determining if a publication achieves the desired impact. Based on this information, we establish the median values of this historical data to calculate the minimum values that a story should achieve for engagement to occur (i.e., .027% for the share rate and .001% for the comment rate).

Third, we calculate an overall index of SMEBs with three values. We assign a value of 0 if a story receives scores lower than .027% for the share rate and lower than .001% for the comment rate; in other words, the SMEB of a story is considered low if it falls below the thresholds established for *both* shares *and* comments. We assign the value of 1 if a story obtains a score equal to or higher than .027% for the share rate, or equal to or higher than

---

[2] We do not account for the number of likes because Instagram did not provide this metric for ephemeral content when we obtained the data.

.001% for the comment rate; in other words, the SMEB is considered medium if it meets or exceeds the threshold for *either* the share rate *or* the comment rate, but not both. We assign the value of 2 if a story obtains a score equal to or higher than .027% for the share rate and equal to or higher than .001% for the comment rate; in other words, the SMEB is considered high if it meets or exceeds the thresholds for *both* the share rate *and* the comment rate.

For *visual elements*, we identify four visual resources applicable to social media: hue, saturation, brightness (HSB), and visual richness. Hue denotes the specific tone of color (Yu, Xie, and Wen 2020). Saturation is the intensity of a hue, ranging from gray to vivid color (Hsieh et al. 2018). Brightness is the amount of lightness (vs. darkness) (Hsieh et al. 2018). Visual richness is the diversity generated by objects in visual content (Zhao et al. 2023). To measure HSB, we use the Clarifai Artificial Intelligence tool, which provides color percentages in hexadecimal code (e.g., #000000 represents black). This code can be directly transformed to indicate values for HSB (Labrecque 2020). We performed this procedure for each pixel of the story to obtain global HSB values. For video-based stories, the global values represent the average of the values of all frames. In comparison to similar tools such as YOLOV2 and Google Cloud Vision, Clarifai provides the most appropriate output labels in a marketing context and enables the identification of multiple content elements such as objects or colors (Nanne et al. 2020). Visual richness was measured by the file size (JPEG), consistent with arguments advanced in prior research (Overgoor et al. 2022; Pieters, Wedel, and Batra 2010; Tuch et al. 2009). File size is considered a reliable, valid, and objective measure in visual environments such as websites or advertisements because an image with a larger file size contains more visual information (Myers et al. 2020; Tuch et al. 2009). Moreover, to focus solely on the visual aspects of video-based stories and mitigate the impact of other elements on the file size (e.g., auditory elements), we used the Photoshop program to convert the original file into individual frames. This process was conducted without altering any visual aspects, enabling us to determine the size of each frame. We then totaled the sizes of these frames to obtain the overall file size. To eliminate potential biases between images and video-based stories, we also measured the size of the former using Photoshop.

We then used dummy conditions to determine whether each of the four visual resources —hue, saturation, brightness, and visual richness— was low (0) or high (1) in a specific story. A story has high hue, saturation, or brightness if Clarifai assigns a score exceeding 50% (Yu, Xie, and Wen 2020); hence, we assign such stories values of 1. Visual richness is considered high (and thus equal to 1) when the size of the image exceeds 200kb, which corresponds to files that fill the screen in most devices, thus signifying a minimum size for rich, complete images. Finally, we formed an overall visual index, ranging from 0 to 4, by summing the values of the four dummy conditions.

We operationalized the *auditory elements* based on the presence of sound (music or speech) and its speed. Sound was considered present if the story contained music or speech. Music implies the inclusion of a song in the story; speech indicates the presence of spoken words. In our data set, those stories containing music or speech do not duplicate the information presented in the text. Regarding the speed of sound, music tempo is the number of beats per minute (BPM) (Oakes and North 2008), and speech rate is the number of words per minute (WPM) (Yang, Yang, and Zhou 2022). Both tempo and speech rates can impact customer responses; faster speeds require more processing resources and increase the risk of information overload (Hahn and Hwang 1999; Krishna 2012). Two authors first identified whether each story contained music or speech. If absent, the story was assigned a value of 0. To assess the speed of sound, we measured tempo using an online metronome. Speed of speech reflects the number of words in the story, which we used to calculate the speech rate, that is, the number of words in the story divided by the duration of the story in minutes. We assigned a value of 1 (i.e., slow tempo or speech rate) for a tempo of less than 130 BPM

(Oakes and North 2008) or a speech rate of less than 180 WPM (Rodero 2016) and a value of 2 (i.e., fast tempo or speech rate) for a tempo equal to or higher than 130 BPM (Oakes and North 2008) or a speech rate equal to or higher than 180 WPM (Rodero 2016). These auditory elements thus enable us to distinguish three levels of intensity: 0 signifies the absence of sound, 1 indicates slow-speed sounds, and 2 represents fast-speed sounds.

For *linguistic elements*, we gauged three types of resources: text, hashtags, and mentions. Hashtags are prefaced by a hash symbol (#); mentions are prefaced by @. Three external human coders with the requisite expertise to evaluate social media content independently coded the presence of text, hashtags, and mentions through the use of # and @. Intercoder reliability was ensured because the value of Krippendorff's alpha (.952) exceeded the threshold established in the literature (.67, in Krippendorff 2004). Any discrepancies between coders were resolved through discussion, and a unique score was assigned to each story. We then used dummy conditions to indicate the absence (0) or presence (1) of each element within the story. From there we constructed a linguistic intensity index ranging from 0 to 2, where 0 denotes the absence of text, hashtags, and mentions; 1 indicates the presence of one resource; and 2 signifies the presence of at least two resources.

In evaluating *symbolic elements*, we considered the inclusion of two resources in stories: emojis and GIFs. Emojis are small, static icons used in interpersonal communication to convey ideas, actions, or emotions (Bashirzadeh, Mai, and Faure 2021; Boutet et al. 2021). GIFs are short, looping sequences of movement that add informativeness, dynamism, and expressiveness (Bashirzadeh, Mai, and Faure 2021; Surucu-Balci, Balci, and Yuen 2020). Since dynamic images increase perceptions of enrichment (Bashirzadeh, Mai, and Faure 2021; Deng et al. 2021), GIFs should convey broader, more intense symbolic meanings than emojis. Three external human coders independently evaluated the symbolic elements and assigned values based on the intensity of the symbols in the story. The scores ranged from 0 to 2, where 0 indicated the absence of symbols, 1 indicated the presence of emojis, and 2 indicated the presence of GIFs. No story in our sample contained both emojis and GIFs. Intercoder reliability was strong (Krippendorff's alpha: .956); any remaining discrepancies were resolved through discussion.

For *social elements*, we assessed the presence of people in the story and their interactions with the focal customer. Interactions can be indirect or direct. If indirect, the people who appear in the story do not establish eye contact or make gestures directed toward the customer (e.g., they appear in a photo looking at the horizon). If direct, the people in the story connect and communicate directly with the customer (e.g., a direct gesture such as a handshake or gaze). The presence or absence of people was initially identified by Clarifai, followed by independent evaluations from three external human coders to determine whether the people in the story interacted directly with the customer. Similar to Dhanesh, Duthler, and Li (2022), we categorized social elements into three levels of intensity: (0) the absence of people in the story; (1) the presence of people in the story with no direct interaction with the customer; and (2) the presence of people directly interacting with the customer. Intercoder reliability was high (Krippendorff's alpha: .987); any remaining discrepancies were resolved through discussion.

To gauge *emotional elements*, we focused on whether each story included emotional content representing emotions, feelings, or moods. Three human coders, unaware of the study's purpose, evaluated the emotional references within each story. Human coders can identify certain characteristics of messages, such as humor, anger, or irony, which automated coding systems may struggle to capture; they also applied homogeneous criteria to examine each publication. After receiving precise instructions, the coders independently assessed whether the story contained references to emotions, feelings, and moods and, if so, the intensity of these emotional references. Specifically, we assigned a value of 0 if the human coders did not identify any reference to emotions, feelings, or moods (e.g., a story featuring a mountain). We assigned a value of 1 if the coders identified an implicit or indirect reference to emotions, feelings, or moods (e.g., a story with a ski lift in the foreground, while a mother gives skiing lessons in the background, might evoke feelings of tenderness and care; however, while present, such sentiments do not constitute the central point of the content). We assigned a value of 2 if the coders identified a clear and easily recognizable reference to emotions, feelings, or moods (e.g., a story depicting a couple engaged in a prominently featured kiss, openly expressing their love). Intercoder reliability was high (Krippendorff's alpha: .748); any remaining discrepancies were resolved through discussion.

Table 2: Conditions and Calibration.

Conditions /outcome	**Resources**	**Levels**	**Fuzzy-set thresholds**
Visual	**Hue:** predominance of colors (vs. black and white): Low hue < 50% High hue ≥ 50% **Saturation:** Low saturation < 50% High saturation ≥ 50% **Brightness:** Low brightness < 50% High brightness ≥ 50% **Visual richness:** Low object < 200kb High object ≥ 200kb	0: Low hue, saturation, brightness, and visual richness 1: One resource with high value: Hue, saturation, brightness, or visual richness 2: Two resources with high value: Hue, saturation, brightness, or visual richness 3: Three resources with high value: Hue, saturation, brightness, or visual richness 4: High hue, saturation, brightness, and visual richness	0 (No membership): Levels 0, 1 .5 (Cross-over point): Level 2 1 (Full membership): Levels 3, 4
Auditory	**Absence of sound** **Presence of slow sound:** Slow tempo of music < 130 BPM Slow speech rate < 180 WPM **Presence of fast sound:** Fast tempo of music ≥ 130 BPM Fast speech rate ≥ 180 WPM	0: Absence of sound 1: Slow sound 2: Fast sound	0 (No membership): Level 0 .5 (Cross-over point): Level 1 1 (Full membership): Level 2
Linguistic	**Absence/presence of text:** Absence of text Presence of text **Absence/presence of hashtags:** Absence of hashtags Presence of hashtags **Absence/presence of mentions:** Absence of mentions	0: Absence of text, hashtags, and mentions 1: Presence of one resource: Text, hashtags, or mentions 2: Presence of at least two resources: Text and/or hashtags and/or mentions	0 (No membership): Level 0 .5 (Cross-over point): Level 1 1 (Full membership): Level 2

	Presence of mentions		
Symbolic	**Absence of symbolic resources** **Presence of symbolic resources:** Emoji(s) GIF(s)	0: Absence of emoji(s) and GIF(s) 1: Presence of emoji(s) 2: Presence of GIF(s)	0 (No membership): Level 0 .5 (Cross-over point): Level 1 1 (Full membership): Level 2
Social	**Absence of people** **Presence of people:** Type of interaction: Indirect interaction Direct interaction	0: Absence of people 1: Presence of people without establishing eye contact or making a gesture directed toward the customer 2: Presence of people with direct connection and communication with the customer (e.g., talking with, or making an explicit gesture)	0 (No membership): Level 0 .5 (Cross-over point): Level 1 1 (Full membership): Level 2
Emotional	**Absence of emotional references** **Presence of emotional references:** Low intensity High intensity	0: Absence of references to emotions, feelings, or moods 1: Presence of indirect or implicit references to emotions, feelings, or moods 2: Presence of clear and central references to emotions, feelings, or moods	0 (No membership): Level 0 .5 (Cross-over point): Level 1 1 (Full membership): Level 2
Outcome: SMEBs	**Low share/comment rate:** Low share ≤ .001% Low comments ≤ .027% **High share/comment rate:** High share > .001% High comment > .027%	0: No SMEBs: Low in share rate and low in comment rate 1: Low SMEBs: High in one category, low in the other category 2: High SMEBs: High in share rate and high in comment rate	0 (No membership): Level 0 .5 (Cross-over point): Level 1 1 (Full membership): Level 2

It is noted that fsQCA does not assume condition independence and is unconcerned with assumptions inherent to linear regressions (e.g., no multicollinearity, no heteroscedasticity, no autocorrelations) (Dusa 2020). Nonetheless, we conclude that the conditions studied here exhibit a high degree of independence as all the correlations are low and even non-significant, being similar to values obtained in other fsQCA studies such as Frösén et al. (2016). Table 3 presents the correlations, means, and standard deviations for the content elements and SMEBs.

Table 3: Correlations, Means, and Standard Deviations of Content Elements and SMEBs.

Content element	1.	2.	3.	4.	5.	6.	7.
1. Visual	1						
2. Auditory	.37**	1					
3. Linguistic	-.13**	.06	1				
4. Symbolic	.03	.02	.03	1			
5. Social	.10*	.29**	.04	.00	1		
6. Emotional	.08	.25**	.15**	.15**	.35**	1	
7. SMEBs	.12**	.08	.00	-.07	-.06	.08	1
M	2.20	.32	.98	.74	.78	2.08	1.19
SD	.70	.60	.61	.93	.66	1.05	.73

* $p < .05$; ** $p < .01$.

Notes: SMEBs: Social media engagement behaviors.

Calibration and Transformation of Conditions

For fsQCA, the studied "conditions and the outcome are [first] calibrated and transformed into fuzzy-set scores, ranging from 0 to 1" (Salonen et al. 2021, p. 150). Specifically, calibration involves determining the thresholds for full membership (1), indifference or the cross-over point (.5), and non-membership (0) for each condition (Salonen et al. 2021). We measured the intensity of each condition, with anchors ranging from the absence of the condition to its full presence. This enabled us to use the high and low values of the conditions to determine full non-membership (0) and full membership (1) as fuzzy-set membership scores. A membership score of .5 denotes an intermediate point, simultaneously a member and a non-member of the fuzzy set, meaning it represents a point of maximum ambiguity (Pappas and Woodside 2021). To avoid difficulties with analyzing cases that scored exactly .5, we added .001 to all conditions scoring below 1 (Fiss 2011). Table 2 contains the fuzzy-set thresholds used for each condition and calibration.

Construction of Truth Table

A truth table gathers all configurations that could occur into a matrix; each row corresponds to a potential combination. By assigning cases to these different configurations, we can determine if they lead to the outcome (Salonen et al. 2021). Similar to Frösén et al. (2016), we used the truth table algorithm provided by the fs/QCA 3.0 software package (Ragin 2008). The resulting truth table consists of 64 configurations. Following fsQCA guidelines, we retained relevant and consistent configurations based on frequency and consistency thresholds (Ragin 2008; Salonen et al. 2021). Frequency "describes how many cases in the sample are explained by a configuration" (Pappas and Woodside 2021, p. 10). Consistency indicates how well a configuration consistently leads to the outcome (Frösén et al. 2016). Given the larger size of our data set, we set the threshold of three as the minimum frequency, as recommended for frequency in samples exceeding 150 cases (Fiss 2011; Frösén et al. 2016; Pappas and Woodside 2021; Ragin 2008). Consistency ranges from 0 to 1; following Rihoux and Ragin (2009), we set a minimum acceptable consistency for configurations at .75. To avoid concurrent subset relations of configurations, we used a proportional reduction in inconsistency score of .7 (Pappas and Woodside 2021). The truth table includes 32 configurations, with 15 observed in the sample, consistent with prior research (e.g., Salonen et al. 2021) (Appendix).

Results

We start by examining whether any condition is necessary for attaining the outcome (Schneider and Wagemann 2010). According to the literature, the minimum values for the necessity of a causal condition are .90 for consistency and .75 for coverage (Ragin 2006; Schneider and Wagemann 2010). Since the consistency scores of the studied conditions are all below .9 (i.e., they range between .27 and .86), none of the conditions emerge as necessary for SMEBs. In other words, none of the social media content elements are necessary for high SMEBs to occur.

For assessing sufficient conditions, fsQCA offers three solutions: complex, parsimonious, and intermediate. Following fsQCA guidelines, we report the intermediate solution (Frösén et al. 2016; Ragin 2008; Salonen et al. 2021), which also indicates core and periphery conditions, as a more accurate view of the findings (Fiss 2011). Core conditions, which appear in the parsimonious and intermediate solutions, exhibit strong causal relationships with the outcome. Peripheral conditions, which appear only in the intermediate solution, show weaker relationships with the outcome; however, they reinforce aspects of the core conditions (Fiss 2011; Salonen et al. 2021).

Table 4 presents the six effective content configurations that consistently lead to SMEBs. We group the configurations by core conditions and present them as four content configurations, some of which include alternative manifestations. The table also includes two evaluative measures: consistency and coverage, both for the overall solution and each individual configuration.

The overall solution consistency measures the extent to which cases (i.e., stories) sharing a particular condition or combination of conditions agree in displaying the examined outcome (Ragin 2006). Our findings achieve a value of .86, surpassing the .80 threshold (Pappas and Woodside 2021); hence, we consider the solution valuable and capable of advancing theory. The overall solution coverage reflects the degree to which the outcome is explained by the configurations as a whole (Ragin 2008; Woodside 2013), akin to the R-squared value (Woodside 2013). Our findings yield an overall solution coverage of .56, above the threshold of .01 proposed in the literature (Ragin 2008; Woodside 2014), representing a substantial proportion of SMEBs. Therefore, we can state that the consistency and coverage of the overall solution meet the minimum criteria for sufficiency (Ragin 2000).

Table 4: Configurations with Sufficient Conditions for SMEBs (Intermediate Solution).

	Loud Content	Informative Content		Affective Content		Relational Content
	C1	C2a	C2b	C3a	C3b	C4
Visual	•	●	●	●	●	
Auditory	●	⊗			⊗	⨂
Linguistic	⨂	●	●	•		●
Symbolic		⨂	⨂			●
Social	•		•	⨂	⨂	●
Emotional	•			●	●	●
Consistency	.98	.88	.90	.88	.87	.89
Raw coverage	.19	.36	.28	.33	.35	.18
Unique coverage	.01	.06	.10	.003	.04	.02
Solution consistency	.86					
Solution coverage	.56					

Notes: Large circles indicate core conditions and small circles indicate peripheral conditions. Black circles indicate the presence of a content element; white circles with "x" indicate the absence of a content element; blank spaces indicate the insignificance of a content element.

The consistency measure for each configuration indicates the degree to which cases sharing similar conditions present the same outcome value (Ragin 2006). Configurations with high consistency scores suggest pathways that almost always lead to the given outcome condition (Elliott 2013). This consistency measure is considered analogous to significance in correlation-based methods (Woodside 2014). Our findings reveal that consistency values for configurations in the intermediate solution range from .87 to .98; thus, we can affirm that all identified configurations are consistently associated with enhanced SMEBs (Pappas and Woodside 2021; Salonen et al. 2021).

The coverage measure for each configuration indicates the relative importance of that configuration in achieving the outcome and is formed by raw and unique coverage (Ragin 2008; Woodside 2013). Raw coverage represents the proportion of the outcome explained by a particular configuration alone (Ragin 2006). As shown in Table 4, the raw coverages of configurations range between .18 and .36, values exceeding those generally accepted in the literature (e.g., Salonen et al. 2021; Woodside 2013; 2014). Unique coverage represents the proportion of the outcome explained exclusively by that configuration, controlling for overlapping explanations by partitioning the raw coverage, “analogous to the partitioning of explained variation in multiple regression” (Ragin 2006, p. 304). The relatively low unique coverage values, particularly for C3a, suggest that the configurations have certain similarities among them rather than representing signifying fully different pathways to the studied outcome. Yet, since the unique coverage values are higher than zero, we conclude that all configurations are relevant for explaining SMEBs (Salonen et al. 2021). According to Ragin (2006, p. 304), assessing both types of coverage “for alternate combinations provides direct evidence of their relative empirical importance.” We identify that informative (C2a) and affective (C3b) content configurations have the best coverage, making them the most meaningful in explaining SMEBs and covering the greatest proportion of cases that can be explained exclusively by them. Overall, the values obtained for consistency, and raw and unique coverage of all configurations closely align with those reported in prior fsQCA studies published in high impact journals (Frösen et al. 2016; Salonen et al. 2021; Woodside 2014). Next, we discuss each identified configuration in more detail. Moreover, we present a *post hoc* analysis that identifies and discusses the content configurations that lead to exceptionally high SMEBs.

Loud Content Configuration

The “loud” content configuration is characterized by the presence of auditory elements and the absence of linguistic elements. The simultaneous inclusion of auditory and linguistic elements with different meanings can reduce message comprehension in advertising (Anand and Sternthal 1990). Yang et al. (2020) also suggest that auditory information may dominate visual-verbal information processing, due to customers’ limited memory resources. However, when auditory elements are combined with reinforcing visual, social, and emotional elements, they can facilitate message perception (Gerdes, Wieser, and Alpers 2014). Visual elements may create a sense of atmosphere and enhance customer immersion (Deng et al. 2021); social elements can enhance the realism of published content, such as when people speak to the camera, creating a sense of face-to-face communication (Wang et al. 2019); and emotional elements can establish a deeper sense of connection (Lee and Theokary 2021).

Examples of loud content configurations from our data set include stories featuring people playing music or talking to customers (see Figure 2). In the first example, a video shows a person playing the saxophone surrounded by people enjoying the experience. The brightness of the image suggests good weather for skiing. In the second example, two people talk directly to the camera; one speaks with enthusiasm, while the other happily explains that their friend is going to start skiing to fulfill a challenge proposed by the firm. The image is dominated by the saturated colors of the friend's skiing equipment.

Figure 2: Examples of Loud Content Configurations.

C1 (music)	**C1 (speech)**
	

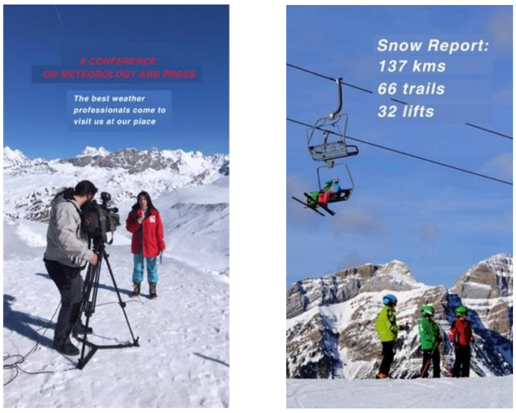
Informative Content Configuration

The "informative" content configuration is characterized by the presence of visual and linguistic elements and the absence of symbolic elements. It emphasizes linguistic elements to educate and inform the audience. However, on social media where attention spans are typically shorter, linguistic content alone may not capture customers' attention; adding visual elements may help (Wang et al. 2019). Dual coding theory (Paivio 1978) suggests a connection between visual and verbal information, which can stimulate each other, leading to fuller comprehension when presented together (Kim and Lennon 2008; Li and Xie 2020; Wang et al. 2019).

Two variations of this configuration encourage SMEBs. Visual and linguistic elements should not be combined with auditory elements (C2a), but they can be combined with social elements (C2b). That is, one way to design informative content is by elaborating stories using visual and linguistic elements without auditory elements; the use of auditory and linguistic elements together when they do not convey the same information may hinder reception of the content (Anand and Sternthal 1990). Notably, as previously explained, this configuration (C2a) achieves the highest coverage, making it the most likely to result in SMEBs. The second path involves complementing visual-linguistic elements with social ones (i.e., the presence of people), making the information easier to understand and more trustworthy (Willis, Palermo, and Burke 2011). Finally, it is important to avoid using symbolic elements such as GIFs, which can be distracting and convey a lack of helpfulness or insufficient employee competence (Huang et al. 2020).

Examples of such content configurations from our data set are shown in Figure 3. The first informative variation (C2a) employs not only diverse visual elements (i.e., a bright landscape and saturated colors) but also includes linguistic elements that provide information and mention the brand. The second informative variation (C2b) combines textual information with bright and saturated images, alongside social elements.

Figure 3: Examples of Informative Content Configurations.

C2a	C2b
	

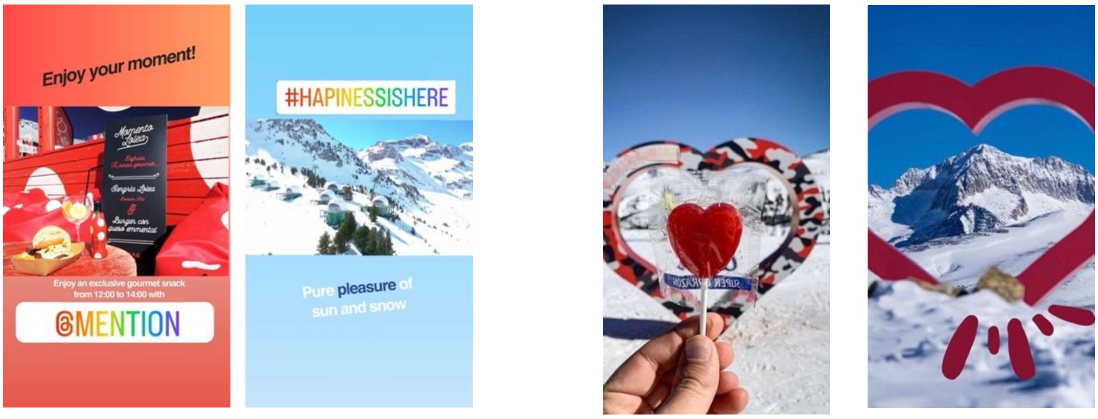
Affective Content Configuration

An "affective" content configuration features the presence of visual and emotional elements while excluding social elements as core conditions. Content that elicits emotions has the potential to create personal connections with customers (Vilches-Montero et al. 2018), and visual elements can also evoke emotions through the inclusion of certain colors (Chi, Pan, and Huang 2021). To design an affective content configuration, firms have two variations: complementing visual-emotional elements with linguistic elements or omitting auditory elements. One approach involves crafting social media content that combines visual, emotional, and linguistic elements. For example, including "#love" in red letters can enhance the emotional function of an image. This aligns with Lee's (2021) assertion that the emotional impact of an image is amplified when the emotional content of the text and the image are considered together. Another option is to rely solely on visual and emotional elements, excluding auditory ones. The inclusion of auditory elements, with their ability to induce emotional reactions, may interfere with the processing of emotional information in other sensory channels (Gerdes, Wieser, and Alpers 2014). Our findings demonstrate that this variation is one of the most effective in achieving SMEBs.

Examples of affective content configurations from our data set are illustrated in Figure 4. The first variation (C3a) depicts colorful stories, with intense colors and emotional text emphasizing enjoyment and happiness. The second variation (C3b) evokes intense emotions without including audio. Both examples include the image of a heart, which conveys tenderness and love, and feature bright, saturated colors.

Figure 4: Examples of Affective Content Configurations.

C3a	C3b



Relational Content Configuration

The "relational" content configuration stands out as the only one marked by the presence of social elements as a core condition; it is also characterized by the inclusion of linguistic, symbolic, and emotional elements and the absence of auditory elements. It thereby exemplifies the importance of selecting a combination of elements that effectively convey complex messages while maintaining coherence. We propose that linguistic, symbolic, social, and emotional elements can be combined to elaborate a complex message, striking an optimal balance that generates SMEBs. Social elements create a feeling of community (Wang et al. 2019), while emotional elements strengthen the customer's connection with the firm (Lee 2021). These relational aspects can be further enhanced with symbolic elements that add playfulness, personality, and novelty (Bashirzadeh et al. 2021; Chi, Pan, and Huang 2021), as well as linguistic elements that clarify information (Kim et al. 2021). This combination of multiple elements aligns with Boutet et al.'s (2021) assertion that emojis provide contextual information that aids in the processing and comprehension of linguistic elements, enhancing personality traits, and magnifying emotional meaning. Similarly, emojis have the potential to amplify emotional intensity through their positivity (e.g., a laughing emoji) or negativity (e.g., a crying emoji) (Das, Wiener, and Kareklas 2019).

Examples of these configurations in our data set reveal people interacting. The first example in Figure 5 depicts closeness by showing two people hugging, accompanied by a GIF designed by the brand and text emphasizing how skiing can unite people. In the second example, three friends are portrayed with a brand-designed GIF and emotive text, indicating that nothing can be better than enjoying skiing with friends. These examples emphasize the relational aspect of the firm by combining linguistic, symbolic, social, and emotional elements.

Figure 5: Examples of Relational Content Configurations.

C4

Post Hoc Analysis. Configurations that Drive Exceptional Social Media Engagement Behaviors

This *post hoc* analysis refines the study to closely examine whether the effective configurations differ in their ability to drive exceptional SMEBs. This study comprised three stages. First, we selected all “success stories” with a score of 2 in SMEBs, totaling 196 cases. Second, we further divided these success stories into below, medium, and large success based on the SMEB. To achieve this, we used the same historical data from other resorts as employed in the original SMEB operationalization, setting the 66th percentile values as thresholds for share and comment rates, instead of using median values. Stories that scored below .051% for the share rate *and* below .018% for the comment rate were assigned a value of 0. As such, a story was considered to be of below success if the SMEBs fell below the thresholds established for both the share and comment rates. Stories scoring equal to or higher than .051% for the share rate *or* equal to or higher than .018% for the comment rate were assigned a value of 1. Thus, the story was considered to be of medium success if the SMEBs were equal to or higher than the thresholds established for *either* the share rate *or* the comment rate, but not both. Stories scoring equal to or higher than .051% for the share rate *and* equal to or higher than .018% for the comment rate were assigned a value of 2. In other words, a story was considered to be of large success if the SMEBs were equal to or higher than the thresholds established for *both* the share rate *and* the comment rate. Third, we identified the membership of cases with below, medium, and large success in the different configurations identified by fsQCA. It should be noted that one case can belong to more than one configuration (Ragin 2008). The absolute numbers of membership for each configuration are shown in Table 5.

We calculated two types of percentages (Ragin 2006). First, we estimated the percentage of cases that were members of a particular configuration (e.g., 13 cases for C1 and large success) divided by the total number of cases at each level of success (e.g., 166 total cases for large success) (see the first percentage in the second line of each cell in Table 5). This percentage shows the importance of each configuration within each level of success. The results show that 47% of the total membership for large success pertains to the affective content configurations (C3). Specifically, C3b is the best-performing configuration, encompassing a quarter of the cases categorized as large success.

Second, we calculated the percentage of cases that belonged to a particular configuration (e.g., 13 cases for C1 and large success) divided by the total number of cases in each configuration (e.g., 37 total cases for C1) (see the second percentage in the third line of each cell in Table 5). This percentage shows the importance of each success level within each configuration, accounting for the effect of the size of the configuration. The results show that 50% of the success stories in the relational content configuration (C4) achieve large success. Thus, although the total number of successful cases for this configuration is low compared to the others, it has the highest probability of achieving large success.

As expected, the findings of the *post hoc* analysis align relatively well with the raw coverage values of each configuration (Table 4).

Table 5: Cross-tabulation of Configurations and Levels of Successful SMEBs.

SMEB success	**Loud Content C1**	**Informative Content C2a**	**Informative Content C2b**	**Affective Content C3a**	**Affective Content C3b**	**Relational Content C4**	**Total number of cases**
Below	10 11% 27%	14 16% 24%	18 21% 20%	17 20% 19%	21 24% 22%	7 8% 22%	87 100% 22%
Medium	14 10% 38%	22 15% 37%	35 24% 40%	33 23% 38%	33 23% 34%	9 6% 28%	146 100% 36%
Large	13 8% 35%	23 14% 39%	35 21% 40%	37 **22%** 42%	42 **25%** 44%	16 10% **50%**	166 100% 42%
Total number of cases	37 9% 100%	59 15% 100%	88 22% 100%	87 22% 100%	96 24% 100%	32 8% 100%	399 100% 100%

Robustness Checks

We used a procedure akin to that adopted by Frösén et al. (2016) to corroborate the fsQCA findings and compare their predictive validity with the results of a regression. Specifically, we assessed the influence of membership configuration score (independent variable) on SMEBs (dependent variable) using ordered logit regression, which is suitable for ordinal dependent variables (Frösén et al. 2016). The membership configuration score reflects the membership score of each story in each of the configurations identified through fsQCA, such that we constructed an independent variable representing the degree to which each story is included "in the configuration where its membership score is highest" (Frösén et al. 2016,

p. 70). The findings confirm that membership in the configurations significantly predicts SMEBs, thus corroborating the fsQCA findings (Model 1, Table 6).

To control for the potential influence of other variables on SMEBs, we included the hour and day of publication (=1 for weekends), transformed into dummy variables. Additionally, to address dynamic endogeneity concerns, because SMEBs can be determined by prior realizations (Hernández-Ortega et al. 2022), we included the lag of SMEBs as a control variable. However, even after controlling for these variables, the configuration membership score continued to significantly predict SMEBs (Model 2, Table 6).

Table 6: Regression Results (Dependent Variable SMEBs).

	Configuration models		**Content element models**	
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Configuration membership score	1.27 (.36)**	1.46 (.38)**		
Visual elements			.30 (.13)*	.32 (.14)*
Auditory elements			.24 (.16)	.19 (.17)
Linguistic elements			.01 (.14)	.06 (.15)
Symbolic elements			-.18 (.09)*	-.22 (.09)*
Social elements			-.55 (.17)**	-.45 (.17)**
Emotional elements			.22 (.09)*	.24 (.09)**
Day of publication		-.44 (.18)**		-.41 (.18)*
Hour dummies	No	Yes	No	Yes
Lag of SMEBs		.45 (.12)**		.40 (.12)**
N	516	516	516	516
Adjusted R-squared	.01	.04	.02	.04

*p < .05; **p < .01. Notes: Coefficients are unstandardized. Numbers in parentheses represent standard errors.

Finally, to evaluate the predictive validity of fsQCA, we conducted additional linear regression analyses with the content elements as individual independent variables, either without (Model 3, Table 6) or with (Model 4, Table 6) control variables. The variance inflation factor values ranged from 1.03 to 1.29, below the threshold of 5, thereby indicating no issues with multicollinearity (Hair et al. 2010). The findings demonstrate a positive effect of visual and emotional elements on SMEBs and a negative effect of symbolic and social elements on SMEBs.

These findings align well with the fsQCA findings, in that visual and emotional elements are core conditions present in at least three configurations. Conversely, symbolic and social elements are core conditions that should be absent in two configurations each. Auditory and linguistic elements demonstrate a more balanced presence versus absence across configurations, which may explain their insignificant effects in the regression analysis (Models 3 and 4, Table 6).

Overall, comparing the fsQCA and regression findings indicates that fsQCA can provide deeper insights into the causal complexity involved in understanding how combinations of content elements predict SMEBs. Although fsQCA can effectively capture causal complexity, the regression model is poorly suited for assessing these types of effects and may nullify or mask more complex effects. As noted in the Method section, the key benefit of fsQCA vis-à-vis regression is the ability to model the asymmetric relationships that often exist in real-world situations. Our findings highlight the value of considering the impact of various elements on SMEBs as a whole, rather than as isolated variables that directly affect SMEBs in a linear way.

Conclusion

In seeking to capture the multimodal communication of social media, we identified various content elements (i.e., visual, auditory, linguistic, symbolic, social, and emotional) and examined how combinations of these elements drive SMEBs. We demonstrate that to explain SMEBs, research should examine content elements as combinations of interlinked semiotic resources rather than isolated elements. With a data set comprising 516 individual Instagram stories and 2,333 frames sourced from a tourism resort operator analyzed with fsQCA, we identify effective content configurations that lead to greater SMEBs: loud, informative, affective, and relational.

Theoretical Contributions

Our study makes two novel contributions to prior literature. First, we advance social media research by identifying key content elements that can be used as semiotic resources in social media (i.e., visual, auditory, linguistic, symbolic, social, and emotional). We thus go beyond the mere visual and linguistic elements, usually addressed in prior research, and provide a more comprehensive perspective of the semiotic resources that can be simultaneously employed in social media as a multimodal means of communication. Furthermore, we challenge the current understanding that tends to focus on the linear effects of individual content elements on customer outcomes (e.g., Pezzuti, Leonhardt, and Warren 2021; Yu and Egger 2021) to consider how firms can use these elements in combination. While it is clearly important to understand the influence of specific content elements on customer outcomes, we must also explore how interactions of diverse content elements drive SMEBs. The findings of this study demonstrate that combinations of content elements, rather than individual elements in isolation, drive outcomes. Whether the presence of an element leads to a positive outcome depends on its fit with other elements.

Second, we identify four "ideal" configurations of content elements that enhance SMEBs. Overall, the findings highlight that there is no single effective approach to designing social media content that prompts SMEBs; rather, several options exist, based on how the different content elements interconnect. The configurations reveal logical patterns and provide rich insights into the effective design of social media content that evokes SMEBs.

Managerial Implications

Given the growing importance of social media in today's digital landscape, understanding how to design social media content to increase SMEBs is imperative for firms that want to remain competitive. This study offers two key managerial implications.

First, social media marketers should consider a range of content elements, beyond visual and linguistic ones, when generating content. Decisions regarding whether to include a song, people, or emotional resources are as important as selecting the colors to include. More importantly, social media marketers should adopt a holistic approach to content design, carefully considering how different content elements fit together rather than seeking to maximize their individual intensity. This highlights the importance of a cohesive strategy for designing content.

Second, our findings reveal four configurations that prompt SMEBs: the loud, the informative, the affective, and the relational. For each content configuration, we provide a "recipe" of the combination of content elements that must be present or absent to encourage SMEBs and that marketers can readily apply. For example, social media marketers opting for an informative content configuration should employ colors, text, hashtags, and mentions to convey firm-related information, while those selecting an affective content configuration should prioritize elements that evoke feelings and emotions to forge connections with the audience.

Limitations and Future Research Avenues

This study identifies content element configurations that drive SMEBs; however, the findings should be interpreted in light of certain limitations. First, the data set was obtained from a single service provider in the tourism sector. While the key argument, that combinations of content elements in social media drive SMEBs, applies to diverse contexts, the focus on tourism may hinder the generalizability of the four specific types of configurations we identify. Future research should thus explore other sectors to determine which configurations lead to high SMEBs across different industries.

Second, we chose ephemeral content as the context of our study because it represents a multimodal form of social media content suitable for our research. While we have theoretically identified content elements applicable to any social media content, ephemeral content also has certain characteristics, such as its fleeting nature, that introduce uncertainty regarding the potential generalization of the four content configurations to other social media content. We suggest that future studies analyze whether these four configurations also drive high SMEBs in other forms of social media content, such as Instagram posts or TikTok videos.

Third, we aimed for a comprehensive yet parsimonious operationalization of six key content elements in social media; however, the range of content elements that could be studied is vast. Future studies could thus investigate more specific aspects to yield even more refined results. For example, future research could separate auditory elements in terms of music and voice, exploring their different effects and analyzing how their joint inclusion with other elements (e.g., voice-over repeating the same information provided in text) affects customers. For social elements, future research could consider not just the presence of people in the post but also their facial expressions, gestures, and types of interactions. Similarly, future research should explore how the storyline of videos (e.g., emotional evolution) influences customer responses (Chowdhury, Olsen, and Pracejus 2008; Grewal et al. 2022).

Finally, while this study focused on SMEBs as an outcome of interest, we do not account for their valence (e.g., negative comments). Therefore, studies might differentiate negative and positive SMEBs, as well as investigate how different content configurations influence other important outcomes for firms, such as website visits, virality, or sales. For ephemeral content-specific outcomes, future research could explore specific relevant outcomes, such as whether customers skip the content or maintain attention (Berger, Moe, and Schweidel 2023).

References

Alamäki, Ari, Juho Pesonen, and Amir Dirin (2019), "Triggering effects of mobile video marketing in nature tourism: Media richness perspective," *Information Processing & Management*, 56 (3), 756–770.

Aleti, Torgeir, Jason I. Pallant, Annamaria Tuan, and Tom van Laer (2019), "Tweeting with the stars: Automated text analysis of the effect of celebrity social media communications on consumer word of mouth," *Journal of Interactive Marketing*, 48, 17–32.

Anand, Punam and Brian Sternthal (1990), "Ease of message processing as a moderator of repetition effects in advertising," *Journal of Marketing Research*, 27 (3), 345–353.

Azer, Jaylan, Thomas Anker, Babak Taheri, and Ross Tinsley (2023), "Consumer-driven racial stigmatization: The moderating role of race in online consumer-to-consumer reviews," *Journal of Business Research*, 157, 113567.

Baker, Julie, A. Parasuraman, Dhruv Grewal, and Glenn B. Voss (2002), "The influence of multiple store environment cues on perceived merchandise value and patronage intentions," *Journal of Marketing*, 66 (2), 120–141.

Barcelos, Renato H., Danilo C. Dantas, and Sylvain Sénécal (2018), "Watch Your Tone: How a Brand's Tone of Voice on Social Media Influences Consumer Responses," *Journal of Interactive Marketing*, 41 (1), 60–80.

Barger, Victor, James W. Peltier, and Don E. Schultz (2016), "Social media and consumer engagement: A review and research agenda," *Journal of Research in Interactive Marketing*, 10 (4), 268–287.

Bashirzadeh, Yashar, Robert Mai, and Corinne Faure (2021), "How rich is too rich? Visual design elements in digital marketing communications," *International Journal of Research in Marketing*, 39 (1), 58-76.

Bauer, Eric and Ron Kohavi (1999), "An empirical comparison of voting classification algorithms: Bagging, boosting, and variants," *Machine Learning*, 36 (1-2), 105–139.

Berger, Jonah, Wendy W. Moe, and David A. Schweidel (2023), "What Holds Attention? Linguistic Drivers of Engagement," *Journal of Marketing*, 87 (5), 793–809.

Bezemer, Jeff and Gunther Kress (2016), *Multimodality, learning and communication: A social semiotic frame*, Routledge: London and New York.

Bitner, Mary J. (1992), "Servicescapes: The impact of physical surroundings on customers and employees," *Journal of Marketing*, 56 (2), 57–71.

Bleier, Alexander, Colleen M. Harmeling, and Robert W. Palmatier (2019), "Creating effective online customer experiences," *Journal of Marketing*, 83 (2), 98–119.

Boutet, Isabelle, Megan LeBlanc, Justin Chamberland, and Charles Collin (2021), "Emojis influence emotional communication, social attributions, and information processing," *Computers in Human Behavior*, 119, 106722.

Brakus, J. Joško, Bernd H. Schmitt, and Lia Zarantonello (2009), "Brand experience: What is it? How is it measured? Does it affect loyalty?" *Journal of Marketing*, 73, 52–68.

Cao, Dongmei, Maureen Meadows, Donna Wong, and Senmao Xia (2021), "Understanding consumers' social media engagement behaviour: An examination of the moderation effect of social media context," *Journal of Business Research*, 122, 835–846.

Ceylan, Gizem, Kristin Diehl, and Wendy Wood (2023), "From Mentally Doing to Actually Doing: A Meta-Analysis of Induced Positive Consumption Simulations," *Journal of Marketing,* 00222429231181071.

Chen, Cheih-Ying, Kun-Huang Huarng, and Vanessa I. González (2022), "How creative cute characters affect purchase intention," *Journal of Business Research*, 142, 211–220.

Chen, Kuan-Ju and Hoi L. Cheung (2019), "Unlocking the power of ephemeral content: The roles of motivations, gratification, need for closure, and engagement," *Computers in Human Behavior*, 97, 67-74.

Chevalier, Judith A. and Dina Mayzlin (2006), "The effect of word of mouth on sales: Online book reviews," *Journal of Marketing Research*, 43 (3), 345–354.

Chi, Maomao, Meiyu Pan, and Rui Huang (2021), "Examining the direct and interaction effects of picture color cues and textual cues related to color on accommodation-sharing platform rental purchase," *International Journal of Hospitality Management*, 99, 103066.

Chowdhury, Rafi M. M. I., G. Douglas Olsen, and John W. Pracejus (2008), "Affective Responses to Images In Print Advertising: Affect Integration in a Simultaneous Presentation Context," *Journal of Advertising*, 37 (3), 7–18.

Das, Gopal, Hillary J. D. Wiener, and Ioannis Kareklas (2019), "To emoji or not to emoji? Examining the influence of emoji on consumer reactions to advertising," *Journal of Business Research*, 96, 147–156.

Demoulin, Nathalie and Kim Willems (2019), "Servicescape irritants and customer satisfaction: The moderating role of shopping motives and involvement," *Journal of Business Research*, 104, 295–306.

Deng, Qi, Yun Wang, Michel Rod, and Shaobo Ji (2021), "Speak to head and heart: The effects of linguistic features on B2B brand engagement on social media," *Industrial Marketing Management*, 99, 1–15.

Dhanesh, Ganga, Gaelle Duthler, and Kang Li (2022), "Social media engagement with organization-generated content: Role of visuals in enhancing public engagement with organizations on Facebook and Instagram," *Public Relations Review,* 48(2), 102174.

Dhaoui, Chedia and Cynthia M. Webster (2021), "Brand and consumer engagement behaviors on Facebook brand pages: Let's have a (positive) conversation," *International Journal of Research in Marketing*, 38 (1), 155–175.

Dolan, Rebecca, John Conduit, John Fahy, and Steve Goodman (2016), "Social media engagement behaviour: A uses and gratifications perspective," *Journal of Strategic Marketing*, 24 (3–4), 261–277.

Dusa, Adrian (2020), *Comparative analysis using Boolean algebra*. SAGE Research Methods Foundations.

Elliott, Thomas (2013), *Fuzzy set qualitative comparative analysis: An introduction*. Research Notes. Statistics Group, UCI.

Fainshmidt, Stav, Michael A. Witt, Ruth V. Aguilera, and Alain Verbeke (2020), "The contributions of qualitative comparative analysis (QCA) to international business research," *Journal of International Business Studies*, 51 (4), 455–466.

Farace, Stefania, Anne Roggeveen, Francisco Villarroel Ordenes, Ko De Ruyter, Martin Wetzels, and Dhruv Grewal (2020), "Patterns in motion: How visual patterns in ads affect product evaluations," *Journal of Advertising*, 49 (1), 3-17.

Filieri, Raffaele, Zhibin Lin, Giovanni Pino, Salma Alguezaui, and Alessandro Inversini (2021), "The role of visual cues in eWOM on consumers' behavioral intention and decisions," *Journal of Business Research,* 135, 663–675.

Fiss, Peer C. (2011), "Building better causal theories: A fuzzy set approach to typologies in organization research," *Academy of Management Journal*, 54 (2), 393–420.

Flewitt, Rosie, Sara Price, Terhi Korkiakangas (2019), "Multimodality: Methodological explorations," *Qualitative Research*, 19 (1), 3-6.

Frösén, Johanna, Jukka Luoma, Matti Jaakkola, Henrikki Tikkanen, and Jaakko Aspara (2016), "What counts versus what can be counted: The complex interplay of market orientation and marketing performance measurement," *Journal of Marketing*, 80 (3), 60–78.

Gerdes, Antje B. M., Matthias J. Wieser, and Georg W. Alpers (2014), "Emotional pictures and sounds: A review of multimodal interactions of emotion cues in multiple domains," *Frontiers in Psychology*, 5, 1351.

Grewal, Dhruv, Dennis Herhausen, Stephan Ludwig, and Francisco Villarroel Ordenes (2022), "The future of digital communication research: Considering dynamics and multimodality," *Journal of Retailing*, 98, 224-240.

Grobelny, Jerzy and Rafal Michalski (2015), "The role of background color, interletter spacing, and font size on preferences in the digital presentation of a product," *Computers in Human Behavior*, 43, 85–100.

Hahn, Minhi and Insuk Hwang (1999), "Effects of tempo and familiarity of background music on message processing in TV advertising: A resource-matching perspective," *Psychology & Marketing*, 16 (8), 659–675.

Hair, Joseph F., W.C. Black, Barry J. Babin, and Rolph E. Anderson (2010), *Advanced diagnostics for multiple regression: A supplement to multivariate data analysis.* Pearson Prentice Hall Publishing.

Hernández-Ortega, Blanca, Michael A. Stanko, Rishika Rishika, Francisco-Jose Molina-Castillo, and José Franco (2022), "Brand-generated social media content and its differential impact on loyalty program members," *Journal of the Academy of Marketing Science*, 50 (5), 1071–1090.

Hernández-Ortega, Blanca, Héctor San Martín, Ángel Herrero, and José L. Franco (2020), "What, how and when? Exploring the influence of firm-generated content on popularity in a tourism destination context," *Journal of Destination Marketing & Management*, 18, 100504.

Holiday, Steven, Jameson L. Hayes, Haseon Park, Yuanwei Lyu, and Yang Zhou (2023), "A multimodal emotion perspective on social media influencer marketing: the effectiveness of influencer emotions, network size, and branding on consumer brand engagement using facial expression and linguistic analysis," *Journal of Interactive Marketing*, 10949968231171104.

Hsieh, Yi-Ching, Hung-Chang Chiu, Yun-Chia Tang, and Monle Lee (2018), "Do colors change realities in online shopping?" *Journal of Interactive Marketing*, 41, 14–27.

Huang, Guei-Hua, Chun-Tuan Chang, Anil Bilgihan, and Fevzi Okumus (2020), "Helpful or harmful? A double-edged sword of emoticons in online review helpfulness," *Tourism Management*, 81, 104135.

HubSpot (2022), "The HubSpot Blog's 2023 social media marketing report: Data from 1200+ global marketers," (accessed May 20, 2023), https://blog.hubspot.com/marketing/hubspot-blog-social-media-marketing-report/.

Jewitt, Carey (2013), *Multimodal methods for researching digital technologies*. The SAGE Handbook of Digital Technology Research, 250–265.

Kim, Molan, Seung M. Lee, Sanghak Choi, and Sang Y. Kim (2021), "Impact of visual information on online consumer review behavior: Evidence from a hotel booking website," *Journal of Retailing and Consumer Services*, 60, 102494.

Kim, Minjeong and Sharron Lennon (2008), "The effects of visual and verbal information on attitudes and purchase intentions in internet shopping," *Psychology & Marketing*, 25 (2), 146–178.

Klostermann, Jan, Anja Plumeyer, Daniel Böger, and Reinhold Decker (2018), "Extracting brand information from social networks: Integrating image, text, and social tagging data," *International Journal of Research in Marketing*, 35 (4), 538–556.

Kress, Gunther R. (2010), *Multimodality: A social semiotic approach to contemporary communication.* Routledge.

Kress, Gunther R. (2011), "Partnerships in research': Multimodality and ethnography," *Qualitative Research*, 11 (3), 239-260.

Krippendorff, Klaus (2004), *Content Analysis: An Introduction to Its Methodology*. SAGE Publications.

Krishna, Aradhna (2012), "An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior," *Journal of Consumer Psychology*, 22 (3), 332–351.

Kumar, Ashish, Ram Bezawada, Rishika Rishika, Ramkumar Janakiraman, and P.K. Kannan (2016), "From social to sale: The effects of firm-generated content in social media on customer behavior," *Journal of Marketing*, 80 (1), 7–25.

Labrecque, Lauren I. (2020), "Color research in marketing: Theoretical and technical considerations for conducting rigorous and impactful color research," *Psychology & Marketing*, 37 (7), 855–863.

Lam, Long W., Ka W Chan., Davis Fong, and Freda Lo (2011), "Does the look matter? The impact of casino servicescape on gaming customer satisfaction, intention to revisit, and desire to stay," *International Journal of Hospitality Management*, 30 (3), 558–567.

Laroche, Michel, Rong Li, Marie-Odile Richard, and Mi Zhou (2022), "An investigation into online atmospherics: The effect of animated images on emotions, cognition, and purchase intentions," *Journal of Retailing and Consumer Services*, 64, 102845.

Lee, Jeffrey K. (2021), "Emotional expressions and brand status," *Journal of Marketing Research*, 58 (6), 1178–1196.

Lee, Michael T. and Carol Theokary (2021), "The superstar social media influencer: Exploiting linguistic style and emotional contagion over content?" *Journal of Business Research*, 132, 860–871.

Li, Xi, Mengze Shi, and Xin (Shane) Wang (2019), "Video mining: Measuring visual information using automatic methods," *International Journal of Research in Marketing*, 36 (2), 216–231.

Li, Yiyi and Ying Xie (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.

Liadeli, Georgia, Francesca Sotgiu, and Peeter W. J. Verlegh (2022), "A meta-analysis of the effects of brands' owned social media on social media engagement and sales," *Journal of Marketing*, 87 (3), 406-427.

Liebrecht, Christine, Christina Tsaousi, and Charlotte van Hooijdonk (2021), "Linguistic elements of conversational human voice in online brand communication: Manipulations and perceptions," *Journal of Business Research*, 132, 124–135.

Lin, Hsin-Chen, Hepsi Swarna, and Patrick F. Bruning (2017), "Taking a global view on brand post popularity: Six social media brand post practices for global markets," *Business Horizons*, 60 (5), 621-633.

Lindner, Thomas, Jonas Puck, and Alain Verbeke (2022), "Beyond addressing multicollinearity: Robust quantitative analysis and machine learning in international business research," *Journal of International Business Studies*, 53 (7), 1307-1314.

Luangrath, Andrea W., Joann Peck, and Victor A. Barger (2017), "Textual paralanguage and its implications for marketing communications," *Journal of Consumer Psychology*, 27 (1), 98-107.

Ludwig, Stephan, Ko de Ruyter, Mike Friedman, Elisabeth C. Brüggen, Martin Wetzels, and Gerard Pfann (2013), "More than words: the influence of affective content and linguistic style matches in online reviews on conversion rates," *Journal of Marketing*, 77 (1), 87–103.

Mahoney, James and Gary Goertz (2006), "A tale of two cultures: contrasting quantitative and qualitative research," *Political Analysis* 14, 227-249.

Mariani, Marcello M., Marco Di Felice, and Matteo Mura (2016), "Facebook as a destination marketing tool: Evidence from Italian regional Destination Management Organizations," *Tourism Management*, 54, 321–343.

McShane, Lindsay, Ethan Pancer, and Maxwell Poole (2019), "The influence of B to B social media message features on brand engagement: A fluency perspective," *Journal of Business-to-Business Marketing*, 26 (1), 1–18.

Meyer, Alan D., Anne S. Tsui, and C.R. Hinings (1993), "Configurational approaches to organizational analysis," *Academy of Management Journal*, 36 (6), 1175–1195.

Mulier, Lana, Hendrik Slabbinck, and Iris Vermeir (2021), "This way up: The effectiveness of mobile vertical video marketing," *Journal of Interactive Marketing*, 55, 1–15.

Myers, Susan D., George D. Deitz, Bruce A. Huhmann, Subhash Jha, and Jennifer H. Tatara (2020), "An eye-tracking study of attention to brand-identifying content and recall of taboo advertising," *Journal of Business Research*, 111, 176–186.

Nanne, Annemarie J., Marjolijn L. Antheunis, Chris G. Van Der Lee, Eric O. Postma, Sander Wubben, and Guda Van Noort (2020), "The use of computer vision to analyze brand-related user generated image content," *Journal of Interactive Marketing*, 50 (1), 156-167.

Oakes, Steve and Adrian North (2008), “Using music to influence cognitive and affective responses in queues of low and high crowd density,” *Journal of Marketing Management*, 24 (5-6), 589-602.

Overgoor, Gijs, William Rand, Willemijn van Dolen, and Masoud Mazloom (2022), “Simplicity is not key: Understanding firm-generated social media images and consumer liking,” *International Journal of Research in Marketing*, 39 (3), 639–655.

Paivio, Allan (1978), “Mental comparisons involving abstract attributes,” *Memory & Cognition*, 6 (3), 199–208.

Pappas, Ilias O. and Arch G. Woodside (2021), “Fuzzy-set qualitative comparative analysis (fsQCA): Guidelines for research practice in information systems and marketing,” *International Journal of Information Management*, 58, 102310.

Pezzuti, Todd, James M. Leonhardt, and Caleb Warren (2021), “Certainty in language increases consumer engagement on social media,” *Journal of Interactive Marketing*, 53, 32–46.

Pieters, Rik, Michel Wedel, and Rajeev Batra (2010), “The Stopping Power of Advertising: Measures and Effects of Visual Complexity,” *Journal of Marketing*, 75 (5), 48–60.

Ponsignon, Frederic, Francois Durrieu, and Tatiana Bouzdine-Chameeva (2017), “Customer experience design: A case study in the cultural sector,” *Journal of Service Management*, 28 (4), 763–787.

Poulsen, Søren V. and Gunhild Kvåle (2018), “Studying social media as semiotic technology: A social semiotic multimodal framework,” *Social Semiotics*, 28 (5), 700-717.

Ragin, Charles C. (2000), *Fuzzy-set social science*. University of Chicago Press.

Ragin, Charles C. (2006), “Set Relations in Social Research: Evaluating Their Consistency and Coverage.” *Political Analysis*, 14 (3), 291–310.

Ragin, Charles C. (2008), *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago Press.

Rietveld, Robert, Willemijn van Dolen, Masoud Mazloom, and Marcel Worring (2020), “What you feel, is what you like: Influence of message appeals on customer engagement on Instagram,” *Journal of Interactive Marketing*, 49, 20–53.

Rihoux, Benoît and Charles C. Ragin (2009), *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*. SAGE.

Robiady, Nurlita D., Nila A. Windasari, and Arfenia Nita (2021), “Customer engagement in online social crowdfunding: The influence of storytelling technique on donation performance,” *International Journal of Research in Marketing*, 38 (2), 492–500.

Rodero, Emma (2016), “Influence of speech rate and information density on recognition: The moderate dynamic mechanism,” *Media Psychology*, 19 (2), 224–242.

Roggeveen, Anne L., Dhruv Grewal, and Elisa B. Schweiger (2020), "The DAST framework for retail atmospherics: The impact of in- and out-of-store retail journey touchpoints on the customer experience," *Journal of Retailing*, 96 (1), 128–137.

Salonen, Anna, Harri Terho, Eva Böhm, Ari Virtanen, and Risto Rajala (2021), "Engaging a product-focused sales force in solution selling: Interplay of individual- and organizational-level conditions," *Journal of the Academy of Marketing Science*, 49 (1), 139–163.

Santini, Fernando O., Wagner J. Ladeira, Diego C. Pinto, Márcia M. Herter, Claudio H. Sampaio, and Barry J. Babin (2020), "Customer engagement in social media: A framework and meta-analysis," *Journal of the Academy of Marketing Science*, 48 (6), 1211-1228.

Schneider, Martin R. and Andreas Eggert (2014), "Embracing complex causality with the QCA method: An invitation," *Journal of Business Market Management*, 7 (1), 312–328.

Schneider, Carsten Q. and Claudius Wagemann (2010), "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets," *Comparative Sociology*, 9 (3), 397–418.

Surucu-Balci, Ebru, Gökcay Balci, and Kum Yuen (2020), "Social media engagement of stakeholders: A decision tree approach in container shipping," *Computers in Industry*, 115, 103152.

Swaminathan, Vanitha, H. Andrew Schwartz, Rowan Menezes, and Shawndra Hill (2022), "The Language of Brands in Social Media: Using Topic Modeling on Social Media Conversations to Drive Brand Strategy," *Journal of Interactive Marketing*, 57 (2), 255–277.

Swani, Kunal, George Milne, and Brian P. Brown (2013), "Spreading the word through likes on Facebook: Evaluating the message strategy effectiveness of Fortune 500 companies," *Journal of Research in Interactive Marketing*, 7 (4), 269–294.

Swani, Kunal, George R. Milne, Brian P. Brown, A. George Assaf, and Naveen Donthu (2017), "What messages to post? Evaluating the popularity of social media communications in business versus consumer markets," *Industrial Marketing Management*, 62, 77–87.

Sykora, Martin, Suzanne Elayan, Ian R. Hodgkinson, Thomas W. Jackson, and Andrew West (2022), "The power of emotions: Leveraging user generated content for customer experience management," *Journal of Business Research*, 144, 997–1006.

Tuch, Alexandre N., Javier A. Bargas-Avila, Klaus Opwis, and Frank H. Wilhelm (2009), "Visual complexity of websites: Effects on users' experience, physiology, performance, and memory," *International Journal of Human-Computer Studies*, 67 (9), 703–715.

van Doorn, Jenny, Katherine N. Lemon, Vikas Mittal, Stephan Nass, Doreén Pick, Peter Pirner, and Peter C. Verhoef (2010), "Customer engagement behavior: Theoretical foundations and research directions," *Journal of Service Research*, 13 (3), 253–266.

Vilches-Montero, Sonia, Hashim Nik M. H. Nik, Ameet Pandit, and Renzo Bravo-Olavarria

(2018), “Using the senses to evaluate aesthetic products at the point of sale: The moderating role of consumers’ goals,” *Journal of Retailing and Consumer Services*, 40, 82–90.

Villarroel Ordenes, Francisco, Dhruv Grewal, Stephan Ludwig, Ko D. Ruyter, Dominik Mahr, and Martin Wetzels (2019), “Cutting through content clutter: How speech and image acts drive consumer sharing of social media brand messages,” *Journal of Consumer Research*, 45 (5), 988-1012.

Wang, Wei, Renee R. Chen, Carol X. Ou, and Steven J. Ren (2019), “Media or message, which is the king in social commerce?: An empirical study of participants’ intention to repost marketing messages on social media,” *Computers in Human Behavior*, 93, 176–191.

Wang, Yunwen (2020), “Humor and camera view on mobile short-form video apps influence user experience and technology-adoption intent, an example of TikTok (DouYin),” *Computers in Human Behavior,* 110, 106373.

Willis, Megan L., Romina Palermo, and Darren Burke (2011), “Social judgments are influenced by both facial expression and direction of eye gaze,” *Social Cognition*, 29 (4), 415–429.

Woodside, Arch G. (2013), “Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory,” *Journal of Business Research*, 66 (4), 463-472.

Woodside, Arch G. (2014), “Embrace•perform•model: Complexity theory, contrarian case analysis, and multiple realities,” *Journal of Business Research,* 67, 2495-2503.

Yang, Cai, Zhi Yang, and Wei Zhou (2022), “Modulating your speech rate: The effect of speech rate on crowdfunding performance,” *Electronic Commerce Research and Applications*, 101211.

Yang, Jialiang, Yaokuang Li, Goran Calic, and Anton Shevchenko (2020), “How multimedia shape crowdfunding outcomes: The overshadowing effect of images and videos on text in campaign information,” *Journal of Business Research*, 117, 6–18.

Yousaf, Anish, Insha Amin, Dhouha Jaziri, and Abhishek Mishra (2021), “Effect of message orientation/vividness on consumer engagement for travel brands on social networking sites,” *Journal of Product & Brand Management,* 30 (1), 44-57.

Yu, Cheung-En, Selina Y. Xie, and Jun Wen (2020), “Coloring the destination: The role of color psychology on Instagram,” *Tourism Management*, 80, 104110.

Yu, Joanne and Roman Egger (2021), “Color and engagement in touristic Instagram pictures: A machine learning approach,” *Annals of Tourism Research*, 89, 103204.

Zhao, Lu, Mingli Zhang, Yaxin Ming, Tao Niu, and Yu Wang (2023), “The effect of image richness on customer engagement: Evidence from Sina Weibo,” *Journal of Business Research*, 154, 113307.

Appendix

Table A: Truth Table.

VIS	AUD	LIN	SYM	SOC	EMO	NUM	SMEBs	RAW	PRI
1	1	**1**	0	1	**0**	8	1	1	1
1	1	**1**	0	0	**1**	6	1	.9588	.8623
1	1	**0**	0	1	**1**	10	1	.9605	.8521
1	1	**0**	1	1	**1**	5	1	.9667	.8505
1	0	**1**	0	1	**0**	28	1	.9470	.8207
1	1	**1**	1	0	**1**	3	1	.9660	.8134
1	1	**1**	0	1	**1**	46	1	.9450	.8129
1	0	**0**	0	0	**1**	4	1	.9103	.7763
1	0	**0**	1	0	**1**	4	1	.9068	.7744
1	0	**1**	0	1	**1**	51	1	.9201	.7739
1	0	**1**	0	0	**1**	25	1	.8917	.7393
1	0	**1**	1	0	**1**	29	1	.8967	.7372
0	0	**1**	1	1	**1**	9	1	.9218	.7246
1	0	**1**	1	0	**1**	42	1	.9034	.7149
1	0	**1**	0	0	**0**	43	1	.8833	.7090
0	0	**1**	0	1	**0**	10	0	.9146	.6972
1	0	**0**	1	1	**1**	11	0	.8925	.6894
1	0	**0**	0	0	**0**	14	0	.8667	.6876
0	0	**1**	0	1	**1**	10	0	.9042	.6845
0	0	**1**	1	0	**1**	9	0	.8891	.6821
0	0	**1**	0	0	**1**	5	0	.8761	.6729
1	0	**0**	0	1	**1**	11	0	.8824	.6669
1	0	**0**	0	1	**0**	15	0	.8823	.6541
0	0	**0**	0	1	**1**	3	0	.8830	.6431
0	0	**0**	1	0	**1**	3	0	.8882	.6392
1	1	**1**	1	1	**1**	35	0	.9151	.6379
0	0	**0**	0	0	**0**	3	0	.8587	.6352
0	0	**1**	0	0	**0**	10	0	.8548	.6178
1	0	**1**	1	0	**0**	12	0	.8539	.5783
0	0	**1**	1	0	**0**	4	0	.8389	.5126
1	0	**0**	1	1	**0**	5	0	.8397	.4695
1	0	**1**	1	1	**0**	24	0	.8516	.4456

Notes: VIS: Visual; AUD: Auditory; LIN: Linguistic; SYM: Symbolic; SOC: Social; EMO: Emotional; SMEBs: Social media engagement behaviors; NUM: number; RAW: Raw consistency; PRI: proportional reduction in inconsistency. In bold: Content elements taken as a basis for the analysis.

We examine the relationship between elements, using those present in a greater number of configurations as the baseline (i.e., visual, linguistic, and emotional). For instance, of the 22 configurations with a "1" in visual, only 32% also have a "1" in auditory, 59% in linguistic, 45% in symbolic, 54% in social, and 64% in emotional. Similarly, of the 20 configurations with a "1" in linguistic, 65% have a "1" in visual, 25% in auditory, 45% in symbolic and social, and 60% in emotional. Finally, of the 20 configurations with a "1" in emotional, 70% have a "1" in visual, 30% in auditory, 60% in linguistic, and 50% in symbolic and social.