Journal of
Research in
Interactive
Marketing

Unlocking the secrets of user engagement: the role of multimodal information and sentiment signals in AI agent design

Shizhen Bai, Yuan Yang, Dingyao Yu and Yongbo Tan
School of Management, Harbin University of Commerce, Harbin, China

Received 24 January 2025
Revised 16 April 2025
4 July 2025
Accepted 6 July 2025

Abstract

Purpose – This study investigates how the information and sentiment of content conveyed by AI agents on digital platforms influence user engagement with multimodal content. Analyzing data on 2,238 agents from the character AI system, we show how text and images activate distinct information-processing channels and jointly shape user interactions.

Design/methodology/approach – We analyze multimodal data from AI agents using structural topic modeling (STM), VADER sentiment analysis, image information entropy and a ResNet-50 deep-learning model. Grounded in dual coding theory (DCT) and the elaboration likelihood model (ELM), we focus on two attributes – informational richness and sentiment polarity – for both text and images, and test their impacts on users' online engagement behavior.

Findings – Text information shows an inverted-U relationship with user engagement. Image visual complexity moderates the effect of textual informational richness. Although aggregate sentiment does not significantly predict engagement, image sentiment amplifies the effect of text sentiment. Images with low emotional intensity create an "emotional vacuum" that increases engagement when paired with positive textual sentiment.

Originality/value – By integrating DCT and ELM, this study offers a new framework for explaining the behavioral effects of multimodal content. It also introduces a method for quantifying the informational and affective attributes of text–image pairs. The findings provide actionable guidance for optimizing digital marketing content and the design of AI-driven conversational agents.

Keywords Multimodal information, User online engagement behavior, AI agent design

Paper type Research article

1. Introduction

AI agents enhance efficiency and reduce costs. According to Grand View Research (2023), the global artificial intelligence market size was valued at USD 279.22 billion in 2024 and is projected to grow at a CAGR of 35.9% from 2025 to 2030. However, their business value hinges on user adoption (Chandra *et al.*, 2022). These NLP-driven virtual assistants (Selvi *et al.*, 2024) handle query-based services (Araujo, 2018), where design elements like text outputs and facial expressions shape user decisions (Ter Stal *et al.*, 2020). Excessive AI-generated information often triggers user indecision, complicating agent selection and increasing time costs (Yankouskaya *et al.*, 2024). Optimizing information disclosure in AI interfaces is thus critical for usability and adoption (Zhou and Wang, 2025).

A primary method of information presentation is the combination of images and text (Eitel and Scheiter, 2015), which simplifies content and mitigates information overload. Existing research has focused on text-image interactions in social media or brand advertising (Dong *et al.*, 2024), examining presentation formats, sentiment expression (Lee and Hong, 2016), and multimodal design (Zhang and Zhang, 2025). However, social media ads primarily attract users through content appeal, where users do not actively choose to interact with specific entities. In contrast, AI agents emphasize functionality and problem-solving (Gizzi *et al.*, 2022), yet existing studies fail to explain user engagement with AI agents adequately. The understanding of the complexity of user-AI agent interaction remains cognitively limited (Kang and Lou, 2022). Interactivity has become an indispensable part of contemporary marketing practices, ranging from retail, services, and communications to live streaming and

Journal of Research in Interactive
Marketing

e-ISSN: 2040-7130
p-ISSN: 2040-7122
DOI 10.1108/JRIM-01-2025-0035

mobile marketing (Wang, 2024). In all these marketing forms, user participation and interaction are critical. To deepen the understanding of how AI agents affect consumer behavior, this study explores how content presentation influences user online engagement behavior (UOEB).

To examine content presentation, we adopt the Elaboration Likelihood Model (ELM), which proposes two cognitive pathways for processing information: central and peripheral (Kitchen *et al*., 2014). We integrate this framework with dual coding theory, highlighting how textual and visual elements act as complementary cognitive channels to facilitate effective information processing (Paivio, 1991). Specifically, research suggests that verbal content with varying specificity levels can effectively trigger both imagery and discourse cognition (Kim, 2019). An important consideration is which cognitive pathway dominates when both are active. Social media content, particularly personal postings, tends to activate users' central cognitive pathway, thereby significantly enhancing UOEB (Schreiner *et al*., 2021). This underscores the necessity of balancing informational or sentiment elements in content design to maximize engagement and influence decision-making (Blasco-Arcas *et al*., 2016).

In defining textual and visual attributes, we apply distinct analytical methodologies. Text attributes are conceptualized into two dimensions: text information, quantified using Structural Topic Modeling (STM) for thematic diversity, and text sentiment, measured through sentiment polarity analysis using the VADER lexical method. Image attributes similarly comprise two dimensions: visual complexity, assessed by image information entropy, and visual sentiment, identified via a ResNet model fine-tuned for sentiment polarity recognition.

Subsequent regression analysis demonstrates that users generally prioritize textual information, with engagement exhibiting an inverted U-shaped pattern, peaking at moderate levels of informational density. High-intensity visual sentiment, even if positive, tends to decrease engagement. Conversely, low-intensity images paired with positively framed text can generate an "emotional vacuum," prompting deeper cognitive engagement.

Grounded in DCT and the ELM, this study investigates how the interaction between textual and visual elements in AI agent communication shapes UOEB. We address four key questions: (1) How does text information displayed by AI agents affect UOEB? (2) Does visual complexity moderate the relationship between text information and UOEB? (3) Does positive text sentiment enhance UOEB? (4) Does visual sentiment moderate the impact of text sentiment on UOEB? In addition, this study provides important theoretical and practical contributions. Based on DCT and ELM, we propose a new way of presenting information in AI agent systems, revealing how the interactions of text and images in AI agents affect UOEB. Combining STM, VADER, image entropy, and ResNet50 models advances multimodal data analysis beyond traditional unimodal approaches. Platform managers should optimize text information volume to reduce overload and enhance interaction. Strategic alignment of text-image sentiment improves content delivery and user engagement.

2. Literature review

2.1 Multimodal data and UOEB in AI platforms

User engagement combines cognitive, emotional, and behavioral participation with brands or service agents (Brodie *et al*., 2011). Hollebeek *et al*. (2014) frame it as a context-dependent, motivation-driven activity, and on AI-driven platforms it manifests as UOEB—likes, comments, shares, and similar actions (Moreno-Munoz *et al*., 2016). In this study, we conceptualize UOEB as the dependent variable. Scholars have moved from self-reports to fine-grained behavioral logs to capture these interactions (Chandra *et al*., 2022). Engagement is shaped by content design, communication strategy, user traits (Hollebeek and Macky, 2019), media richness (Santos *et al*., 2023), and perceived interactivity (Bozkurt *et al*., 2025), which spurs value co-creation (Gao *et al*., 2023) and is reflected in commenting and sharing (Sussman *et al*., 2023). High image quality lifts consumer engagement (Ryu, 2024), while real-time dynamic comments amplify immediate responses in livestreams (Wei, 2022; Zhang *et al*., 2024). Anthropomorphic cues deepen emotional bonds and dependency (Yim *et al*., 2024), whereas ethical design boosts trust and compliance (Morosan and Dursun-Cengizci, 2024).

Multimodal stimuli—such as text, images, video, and audio—provide users with enriched contextual information and enhanced emotional resonance (Riaz, 2019). Among these, attributes of text-image combinations, including sentiment valence, visual complexity, and information volume, have received considerable scholarly attention. However, the existing literature lacks systematic theoretical elaboration and empirical validation regarding the joint effects of these multimodal attributes (Henlein *et al*., 2024). Moreover, while informativeness has been shown to increase click-through rates in social advertisements (Mir and Salo, 2025), and social signals have been found to enhance users' visitation intentions in virtual environments (Sun *et al*., 2023), such research on multimodal information and user responses has predominantly focused on social media or virtual communities, with limited attention to user engagement mechanisms in AI-mediated contexts. Notably, prior studies have indicated that transparent and detailed multimodal information can significantly enhance user trust in AI systems (Kim *et al*., 2021), further underscoring the critical role of multimodal content in AI-driven interaction scenarios. Nevertheless, there remains a significant gap in understanding how multimodal factors collectively influence UOEB in AI-enabled environments. To address this theoretical gap, the present study investigates how multimodal attributes jointly shape UOEB in AI-mediated contexts, thereby providing theoretical insights and practical implications for the development of more effective interactive marketing strategies in AI-driven settings.

2.2 Theoretical framework

This study builds an integrated theoretical framework drawing on Dual Coding Theory (DCT) and the Elaboration Likelihood Model (ELM) to examine how users process informational and sentiment cues in AI-generated multimodal content. DCT posits that individuals process text and image information through parallel yet interconnected cognitive channels, with meaningful alignment between modalities enhancing comprehension and recall (Scheiter *et al*., 2014). ELM explains how individuals engage with messages through two distinct persuasion routes: the "central route", guided by deliberate, content-based elaboration, and the "peripheral route", influenced by affective and contextual cues (Petty and Cacioppo, 1984).

The dual-channel structure of DCT provides the modality through which central and peripheral cues—outlined in ELM—are delivered and integrated. In multimodal context, text and image stimuli rarely function in isolation; instead, they co-construct meaning, with emotional resonance and logical content unfolding simultaneously. This interplay is particularly salient in human-AI interaction, where text and image often merge instrumental functionality with implicit sentiment or information cues.

Recent studies underscore the relevance of this theoretical integration. For example, image-text congruence on social media has been shown to capture attention and enhance message effectiveness (Bigne *et al*., 2021), while credibility assessments often rely more heavily on central-route processing, particularly when logical coherence is salient (Cheung *et al*., 2012). Sentiment expressions in imagery—such as facial cues—convey social intent (Van Kleef, 2009), yet overly intense emotional displays may backfire by reducing perceived competence (Wang *et al*., 2017). Conversely, images with "low sentiment intensity" exert less pressure on the viewer, supporting more neutral, less emotionally saturated interactions.

Within this integrated model, we propose the notion of an "emotional vacuum"—a multimodal communicative state characterized by the relative absence of emotionally charged signals. Unlike high-arousal content that can trigger strong affective responses and biased cognitive processing, an emotional vacuum offers a low-intensity affective environment that allows users to engage more evenly with both informational and relational aspects of the content.

2.3 Research Model and hypothesis development

Building on the integrated framework derived from DCT and ELM, we examine how text and image cues interact to shape UOEB. According to DCT, text is processed through both the central route, focused on cognitive analysis and the peripheral route, driven by sentiment responses. ELM highlights the role of both cognitive and sentimental pathways in persuasion. In this context, text information, referring to the number of textual themes, influences engagement through the central route, while text sentiment, the polarity of textual sentiment, affects engagement via the peripheral route. Image elements such as visual complexity, the density and detail of images, and visual sentiment, the sentiment tone in images, further moderate these relationships. Based on this theoretical grounding, we propose the following hypotheses to explore the impact of these factors on UOEB.

Text material conveys product functionalities as users' primary information source. Studies note that differing information volumes in online vs. offline ads alter initial ad processing, subsequently affecting users' memory and reprocessing during repeated exposure (Chatterjee, 2012). However, multi-theme texts demand higher cognitive effort, and excessive information triggers overload, impairing processing efficiency (Huang *et al*., 2021). Moderate volume optimizes cognitive activation and engagement, while overload depletes cognitive resources, reducing participation. Thus, we propose the following hypothesis:

H1. There is an inverted U-shaped relationship between text information and UOEB.

Text conveys both agent information and sentimental cues. According to the ELM, users process information through either central routes (logic-driven analysis) or peripheral routes (emotional analysis). Positive sentiment textual consistently fosters user identification and engagement (Arapakis *et al*., 2014), whereas negative sentiment textual— even when coupled with moderate information—diminishes interaction, underscoring the asymmetric role of sentiment in shaping engagement outcomes. Therefore, we propose the following hypothesis:

H2. Positive textual sentiment positively influences UOEB

Visual complexity, defined as the density and detail of visual elements (Airoldi and Bischof, 2016), plays a critical role in shaping consumer perceptions and behaviors (Flores *et al*., 2014). Complex visuals interact with cognition through mechanisms like facial recognition and image-text congruence, which enhance engagement, especially on platforms like social media (Li and Xie, 2020). Images provide immediate and intuitive perceptual impact, while text offers deeper semantic information. The synergy between these two elements enhances user identification with virtual characters, thereby increasing user engagement. We argue that visual complexity moderates the relationship between text information and UOEB. Specifically, the presence of images alleviates the cognitive burden of text, preventing overload. Consequently, the typical inverted U-shaped curve between textual information and UOEB flattens, as users engage more effectively with both content types, enhancing overall interaction.

H3. Visual complexity moderates the inverted U-shaped relationship between text information and UOEB, flattening the curve.

Visual content conveys both information and emotional cues. Facial expressions and non-facial emoticons influence customer engagement by triggering interactive behaviors and enhancing participation (Valenzuela-Gálvez *et al*., 2022), with facial expressions functioning as powerful, natural signals of emotional intent (Gross, 2010). However, excessive emotional displays—such as exaggerated happiness in live-stream sales—can disrupt information delivery and reduce performance (Bharadwaj *et al*., 2022). This tension underscores the importance of balance: low-intensity visual emotions, when paired with positive textual sentiment, create an "emotional vacuum" that shifts user attention from affective reactions to cognitive processing. This focused cognitive engagement improves comprehension and increases participation, suggesting that restrained visual sentiment optimizes the interaction between emotional and informational cues. Based on this reasoning, we propose the following hypothesis:

H4. Low visual sentiment creates an "emotional vacuum" that increases user engagement when combined with positively text sentiment.

Building on these hypotheses, the integrated theoretical framework derived from DCT and ELM provides a comprehensive understanding of how textual and visual elements interact in shaping UOEB. To further explore this, we now present a framework diagram (Figure 1) illustrating the relationships among these variables.

3. Methodology

3.1 Sample and data collection

This study analyzed data from Character AI, a leading AI-driven virtual character platform, to assess how content information and sentiment impact user engagement. The system's interface is shown in Figure 2. Data from 2,382 unique AI agents across 24 domains (e.g. entertainment, gaming, education) was collected via Python web scraping. Each agent's profile included name, avatar, domain, chat frequency, likes, and textual descriptions (greeting, personality and

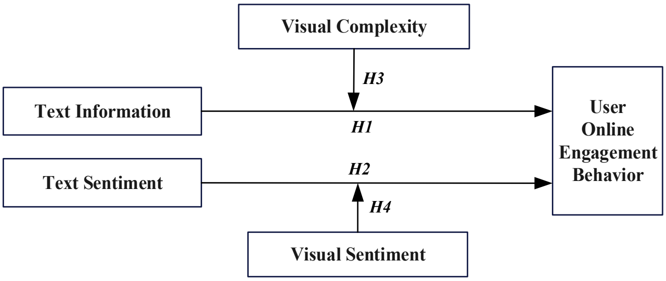

Figure 1. Research model. Source: Authors own work

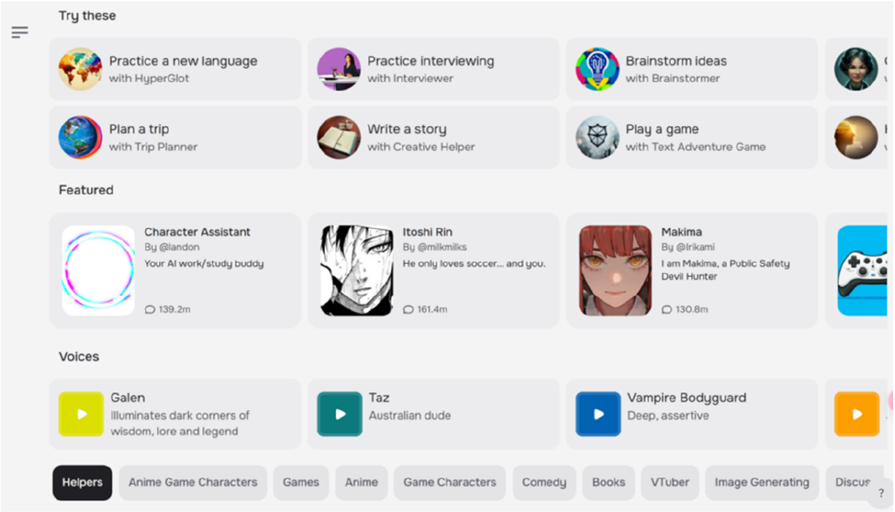

Figure 2. Character AI system page view. Source: https://character.ai/

hobbies). Duplicates and misclassified agents were removed to ensure dataset integrity. Visual complexity and sentiment were evaluated using agent avatars, while textual content quantified information volume and sentiment polarity. Engagement metrics served as dependent variables.

3.2 Text information

The value of information lies in reducing uncertainty. The variable text information captures the diverse information conveyed by AI agents, reflecting the text's ability to fulfill consumers' informational needs. Since we cannot access the specific content of user-AI agent interactions, we measure text information using the static textual disclosed by each AI agent (e.g. profile signatures, greetings, role descriptions, personal backgrounds, hobbies) to assess its impact on user online engagement behavior. To quantify text information, we employ Structural Topic Modeling (STM) (Roberts *et al.*, 2019), where text information is defined as the number of topics identified in the textual content disclosed by each AI agent. As STM is an unsupervised machine learning text-mining technique, we manually determine the optimal number of topics (K). Using the searchK function from the R programming language's STM package, we test models trained on sparse matrices across multiple ranges in parallel. Given the small sample size in this study, we select the optimal K between 5 and 12, as illustrated in Figure 3.

The optimal topic count (K = 8) was selected using the searchK function in R's STM package, balancing four metrics: maximizing semantic coherence and held-out likelihood, while minimizing residuals and lower bound values (Airoldi and Bischof, 2016; Roberts *et al.*, 2019). Topics with low relevance (proportion <5%) were excluded as noise, retaining only significant themes (e.g. high-frequency words in Table 1). The total count of retained topics (>5%) quantified text information, reflecting textual richness and reducing uncertainty.

3.3 Text sentiment

This study applied the VADER method (Hutto and Gilbert, 2014), a lexicon-based tool tailored for social media texts (e.g. tweets, comments). VADER evaluates sentiment using a predefined lexicon that includes emotional terms, contextual modifiers (e.g. punctuation like "!!!",

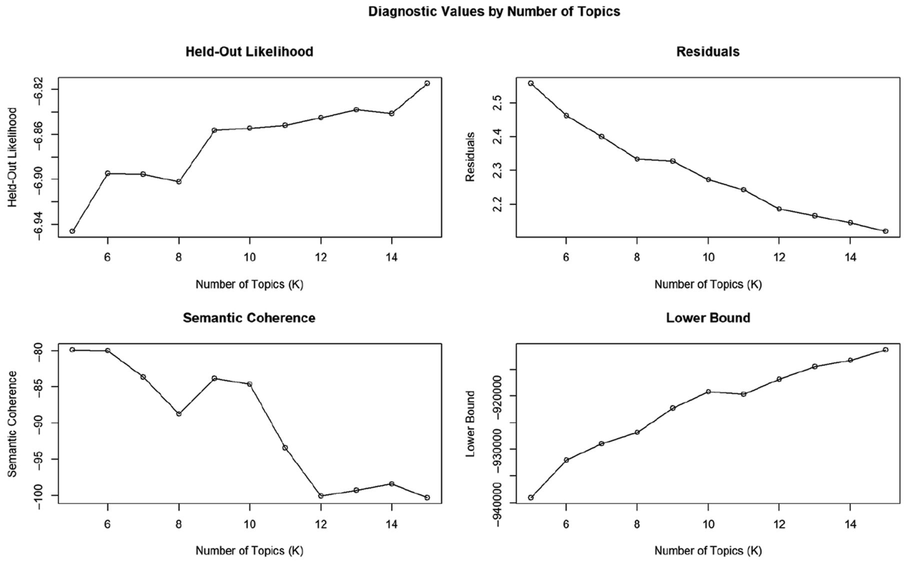

Figure 3. Results of indicators for the assessment of the number of themes. Source: Authors own work

Table 1. Summary results of thematic high-frequency words

Variable	Highest prob
Topic 1 Education	student, teacher, class, histori, studi, school, teach, histor, philosophi, strict
Topic 2 Fiction	human, demon, kill, name, face, manipul, magic, surviv, mani, vampir
Topic 3 Creation, help	charact, help, stori, creat, peopl, will, world, write, explor, provid
Topic 4 Personal character	hair, girl, school, friend, black, love, like, shes, tall, brown
Topic 5 Hobbies	game, love, book, share, play, excel, geek, also, enjoy, read
Topic 6 Diversion	music, movi, actor, famous, interview, play, love, singer, band, show
Topic 7 Personal experience	user, look, just, time, like, back, friend, room, feel, want
Topic 8 Family, responsibilities	love, care, famili, also, person, cold, polit, despit, protect, make

Source(s): Authors' own work

capitalization), and social media-specific expressions. The Compound Score (-1 = negative, $+1$ = positive) synthesizes sentiment intensity and polarity. For ease of processing and alignment with image data, this study categorized text sentiment as follows: texts with scores less than or equal to -0.05 were defined as negative and assigned a value of -1, while texts with scores greater than -0.05 were defined as positive and assigned a value of 1 (van der Waal, 2024).

3.4 User online engagement behavior

In this study, we collected the number of chats and likes between users and AI agents (Pan *et al.*, 2024) and used them as key metrics to measure UOEB. UOEB is a complex behavioral construct reflecting multidimensional psychological and behavioral characteristics in human-AI interactions. Specifically, chats and likes frequency was quantified by counting the number of independent sessions in which users accessed an AI agent's chat interface, as recorded in platform backend logs. Both metrics exhibited severe right-skewed distributions (Shapiro-Wilk: chats skewness = 8.532, likes = 8.133, $p < 0.001$). To address this skewness, a natural log transformation ($X' = \ln(X + 1)$) was applied, which reduced the skewness to -0.214 for likes and -0.558 for chats. Given the high correlation between variables ($r = 0.892^{**}$, $p < 0.01$), Principal Component Analysis (PCA) was used to consolidate chats and likes into a single UOEB index. This approach mitigated multicollinearity while preserving the dimensionality of engagement. The suitability of PCA was confirmed by the Kaiser-Meyer-Olkin (KMO = 0.732) and Bartlett's sphericity tests ($\chi^2 = 317.84$, $p < 0.001$). Eigenvalue decomposition retained the first principal component (eigenvalue = 1.821), explaining 91.1% variance, per Kaiser criterion (eigenvalue >1) and scree test. Loadings for both variables (likes/chats) were 0.954, demonstrating strong explanatory power (>95%) for the synthesized UOEB index (Abdi and Williams, 2010).

3.5 Visual complexity

During interactions with AI agents, the complexity of images plays a critical role in UOEB. Image entropy—which quantifies complexity in terms of detail, variation, and uncertainty (Bruce and Tsotsos, 2009)—is therefore a key metric for optimizing AI agents design. Research shows that higher entropy shifts users' temporal focus and, in turn, shapes consumer decision-making (Biliciler *et al.*, 2022). In this study, we calculate entropy from color histograms (Rigau *et al.*, 2005) to gauge visual impact and information efficiency (Figure 4), aligning our approach with findings on visual attention and emotional response in AI-mediated interactions.

To process the images, we first convert them into grayscale. The grayscale image is then analyzed by calculating the probability distribution of each gray level in the histogram (Equation (1)), where Total Pixels represents the total number of pixels in the image. The probability distribution of pixel values in the grayscale image is given by Equation (2),

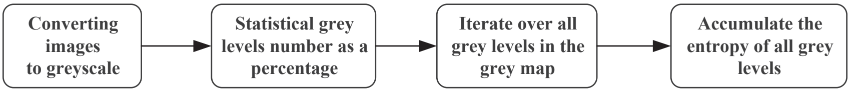

Figure 4. Steps to calculate the entropy of image information. Source: Authors own work

where P(i) denotes the probability of gray level i occurring in the image. Finally, the information content in the image is transformed into an entropy value, which is used as the measure of the information quantity contained in the image.

$$P(i) = \frac{N(i)}{\text{Total Pixels}} \tag{1}$$

$$H = -\sum_{i=0}^{255} p(i)\log_2 p(i) \tag{2}$$

3.6 Visual sentiment

The dataset includes real individuals, movie characters, and primarily anime characters. Since existing models were ineffective for our dataset, we chose ResNet50 for training. The residual structure of ResNet50 (He *et al.*, 2016) effectively captures the stylistic features of anime images and performs well in facial recognition tasks (Li and Lima, 2021). To classify the data, we referenced Face++'s "smile index" (An and Weber, 2021) which quantifies visual sentiment on a 0–100 scale, binarized at 50 (>50 = high, <50 = low). To ensure data accuracy, we recruited 10 computer science graduate students as coders. After standardized training, they annotated 420 images (50 images per coder), resolving discrepancies through consensus. The Pearson correlation (Wang *et al.*, 2020) between human annotations and Face++ predictions was 0.759 ($p < 0.001$), and the inter-rater ICC was 0.860, confirming the reliability of the annotations.

In model training, we fine-tuned ResNet50, pre-trained on ImageNet. To preserve general visual features, we froze the initial convolutional layers and trained the later layers on an augmented dataset (including rotations, flips, scaling, and color jitter). Grayscale conversion further reduced redundancy. The model achieved 83.37% cross-validation accuracy (training loss = 0.39, validation loss = 0.43) and 78% accuracy on 245 test images, demonstrating strong sentiment classification performance. The trained model assigns sentiment scores to AI agent images, as illustrated in Figure 5 (1 = high [left], −1 = low [right]), achieving a balance between specificity and generalization.

Figure 5. Diagram of coding results. Source: https://character.ai/

3.7 Control variables

We posit that longer text may contain more topics, and the number of sentences can reflect the semantic distribution of the text. Additionally, the displayed chat frequency may influence user engagement behavior, as users are more likely to interact with AI Agents that have higher chat counts. Furthermore, we retained the original domain classification of the AI Agents. Therefore, we controlled for text length (Huo *et al.*, 2024), chat frequency (Pittman and Haley, 2023), and domain.

3.8 Descriptive statistics and correlation analysis

To eliminate dimensional differences between variables, we standardized all variables to improve the effectiveness of subsequent analyses and modeling. The descriptive statistics for the variables are shown in Table 2.

To ensure the validity of the regression analysis, we tested the assumptions of the regression model. Residuals met normality, independence, and homoscedasticity. Multicollinearity was absent (VIFs <10). Pearson correlations indicated significant predictor-moderator relationships (Table 3), while control variables showed no intercorrelations. All assumptions were satisfied, enabling robust regression analysis.

Table 2. Descriptive statistics of variables

		Mean	Standard Deviation	Minimum	Maximum	*N*
Dependent variable	Text information	0.00	1.00	−2.11	5.22	2,382
	Text sentiment	0.00	0.803	−1	1	2,382
Independent variable	UOEB	0.00	1.00	−2.39	2.44	2,382
Moderating variable	Visual Complexity	0.00	1.00	−5.84	1.19	2,382
	Visual Sentiment	0.01	1.00	−1	1	2,382
Control variable	Text Length	12.67	8.88	0	151	2,382
	Chats number	0.75	0.433	0	1	2,382
	Type	13.75	7.6	1	24	2,382

Source(s): Authors’ own work

Table 3. The results of correlation analysis

	Text information	Text sentiment	UOEB	Visual complexity	Visual sentiment	Text length	Chats number	Type
text information	1							
text sentiment	0.003	1						
UOEB	0.024	0.002	1					
Visual Complexity	0.011	0.032	−0.022	1				
Visual Sentiment	−0.001	0.007	−0.004	0.027	1			
Text Length	0.013	−0.010	0.000	−0.019	0.036	1		
Chats number	0.032	−0.001	0.198	0.026	−0.002	0.055^{*}	1	
Type	0.002	0.006	-0.095^{**}	−0.005	0.120^{**}	−0.019	−0.010	1

Source(s): Authors’ own work

4. Result

4.1 Baseline results

We used SPSS and Python to perform multiple linear regression analyses to examine the relationship between text information, text sentiment, and UOEB. In Model 1 (see Table 4), we applied the following regression equation (Equation 3), where UOEB was set as the dependent variable, and text information was the independent variable. Control variables were not included in this model. β_0 is the intercept term, β_1 represents the linear effect of text information on UOEB, and β_2 represents the nonlinear effect of text information^2 on UOEB. In Model 2 (Equation 4), we included three control variables. β_3 is the intercept term, β_4 represents the linear effect of text information on UOEB, and β_5 represents the nonlinear effect of text information^2 on UOEB. Additionally, $\delta_1 \delta_2$ capture the effects of text length and chats number on UOEB, respectively, while γ_i represent the regression coefficients for the categorical domain variables.

$$UOEB = \beta_0 + \beta_1 \text{Text Information} + \beta_2 \text{Text Information}^2 + \varepsilon_1 \tag{3}$$

$$UOEB = \beta_3 + \beta_4 \text{Text Information} + \beta_5 \text{Text Information}^2 + \delta_1 \mathit{Text\ Length} + \delta_2 \mathit{Chat\ Number} + \sum_{i=1}^{24} \gamma_i \mathit{Type}_i + \varepsilon_2 \tag{4}$$

The test for the inverted U-shaped curve includes the following three criteria (Haans *et al.*, 2016). First, the results indicate that text information^2 has a significant impact on sales, both without control variables ($\beta_2 = -0.692$, $p < 0.01$) and with control variables ($\beta_5 = -0.362$, $p < 0.001$). Additionally, the quadratic coefficient is significant and negative ($\beta_5 < 0$). Second, the slopes at both ends of the text information range are sufficiently steep. At the low end of text information, the slope is significantly positive (slope $= 1.82 > 0$), whereas at the high end, the slope is significantly negative (slope $= -1.80 < 0$). Third, the turning point of the curve (TP $= 0.514$) lies within the range of text information (see Figure 6). Therefore, a clear inverted U-shaped relationship exists between text information and UOEB.

In addition, regression results from Model 1 show that text length has no significant effect on UOEB, whereas chat frequency and AI agent type significantly influence engagement. These findings support H1. When the volume of information is too low, users may perceive it as insufficient, reducing engagement. Conversely, excessive information leads to cognitive overload, also suppressing engagement. Engagement peaks when information volume reaches a moderate level (e.g. approximately four units), achieving a balance between informativeness and cognitive processing demands. Finally, we rule out the possibility of an S-shaped curve and examine the effect of subject share in the Appendix.

Table 4. The regression results of Models 1

	Without control(Models 1)					
Variable	Beta	*t*	95% CI	Beta	*t*	95% CI
Text information	0.372^{***}	5.958	−0.186–0.369	0.716^{*}	8.147	0.405–0.662
Text information^2	-0.362^{***}	−5.801	−0.044–−0.022	-0.692^{***}	−7.875	−0.079–−0.048
Text length	0.025	1.753	−0.004–0.009			
Chats number	0.703^{***}	49.090	1.555–1.685			
Type	-0.065^{***}	−4.564	−0.012–−0.005			
R^2	0.519			0.026		
F	514.656^{***}			33.217^{***}		

Note(s): *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Source(s): Authors' own work

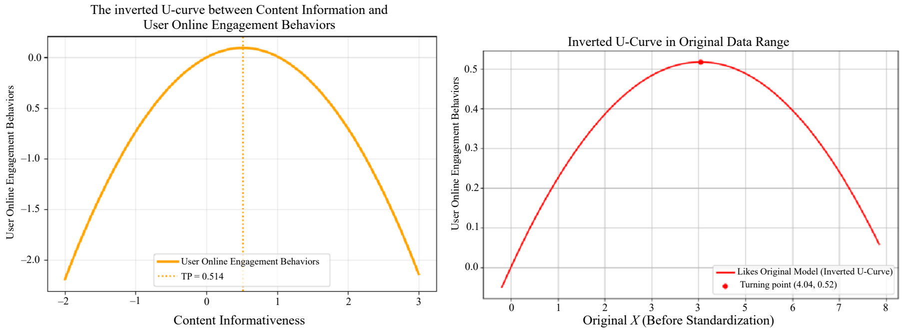

Figure 6. The inverted U-curve between Content Information and User online engagement behavior. Source: Authors own work

As shown in Table 5, Model 3 (Equation 5) examines the relationship between text sentiment and UOEB.

$$UOEB = \beta_6 + \beta_7 \text{Text Sentiment} + \delta_3 Text\ Length + \delta_4 Chat\ Number + \sum_{i=1}^{24} \gamma_i Type_i + \varepsilon_3 \quad (5)$$

The results show that the effect of text sentiment on UOEB is not significant ($\beta_7 = 0.106$, $p > 0.05$), suggesting that its independent effect on UOEB is weak, and therefore H2 is rejected. In the Appendix, we conduct a simple experiment to explore the reasons for rejecting H2 and to confirm the roles of text information and sentiment in the fine-grained probabilistic model for the different processing pathways. The results show that while text sentiment also affects UOEB, the coefficient for information is significantly higher than the coefficient for sentiment. This is consistent with the ELM's "central path" theory, which suggests that users prioritize the logic and utility of information when processing it, and that the actual text of information outweighs its sentimental value.

4.2 Channel analysis

We used SPSS and Python to perform multiple linear regression analyses to examine the relationships among text information, text sentiment, UOEB, visual complexity, and visual sentiment.

As shown in Table 6, Model 4 (Equation 6) tests the moderating effect of visual complexity on the main effect of text information on UOEB.

Table 5. The regression results of Models 3

Variable	Beta	t	95% CI
Text sentiment	0.106	1.974	−0.005–0.212
Text length	0.025	1.757	0.000–0.000
Chat number	0.711^{***}	49.877	1.576–1.705
Type	-0.079^{***}	−3.902	−0.016–−0.005
R^2	0.516		
F	424.235^{**}		

Note(s): *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Source(s): Authors' own work

Table 6. The regression results of Models 4 and Models 5

	Models 5					
Variable	Beta	*t*	95% CI	Beta	*t*	95% CI
Text information	0.250^{***}	3.010	0.098–0.746			
Text information^2	-0.290^{*}	−4.650	−0.485–−0.215			
Visual complexity	-0.220^{*}	−3.858	−0.332–−0.108			
Text sentiment				0.072	−2.008	−0.003–0.214
Visual sentiment				-1.008^{**}	−2.907	−1.687–−0.328
Text information × Visual complexity	-0.090^{***}	−3.833	−0.140–−0.044			
Text information^2 × Visual complexity	0.120^{***}	3.876	0.060–0.180			
Text sentiment × Visual sentiment				-0.035^{*}	−2.479	−0.158–−0.018
Text length	0.049	1.633	0.000–0.000	0.026	1.822	0.000–0.000
Chat number	1.620^{***}	49.065	0.215–0.485	0.711^{***}	49.848	1.576–1.705
Type	-0.008^{***}	−4.405	−0.012–−0.005	-0.060^{***}	−4.203	−0.012–−0.004
R^2	0.518			0.516		
F	321.418^{***}			318.574^{***}		

Note(s): $*p < 0.05$, $**p < 0.01$, $***p < 0.001$
Source(s): Authors' own work

$$\begin{aligned}\text{UOEB} = {} & \beta_8 + \beta_9 \text{Text information} + \beta_{10} \text{Text information}^2 + \beta_{11} \text{Visual complexity} \\ & + \beta_{12} (\text{Text information} \times \text{Visual complexity}) \\ & + \beta_{13} \left(\text{Text information}^2 \times \text{Visual complexity}\right) + \delta_5 \text{Text length} + \delta_6 \text{Chat number} \\ & + \sum_{i-1}^{24} \gamma_i \text{Type}_i + \varepsilon_4 \end{aligned} \tag{6}$$

The analysis shows that visual complexity has a significant negative main effect on UOEB ($\beta_{11} = -0.022$, $p < 0.05$). The interaction between text information and visual complexity is also negative and significant ($\beta_{12} = -0.09$, $p < 0.05$). Notably, the sign of this interaction coefficient ($\beta_{12} < 0$) is opposite to that of the linear term for text information ($\beta_9 = 0.25$), while the coefficient for the quadratic interaction (text information^2 × visual complexity, $\beta_{13} = 0.12$) is opposite to that of the quadratic main term (text information^2, $\beta_{10} = -0.29$).

In Model 4, since $\beta_{13} > 0$, the negative slope between text information and UOEB becomes less steep, flattening the inverted-U curve and shifting its peak leftward (see Figure 7). These results confirm that visual complexity moderates the inverted U-shaped relationship between text information and UOEB, supporting H3. After accounting for visual complexity, the peak of the curve lowers, the decline on both sides slows, and the optimal range of information volume widens. Images supply supplementary cues, reducing users' dependence on text information alone. Consequently, users tolerate a broader range of text volume before negative effects emerge, enhancing system robustness.

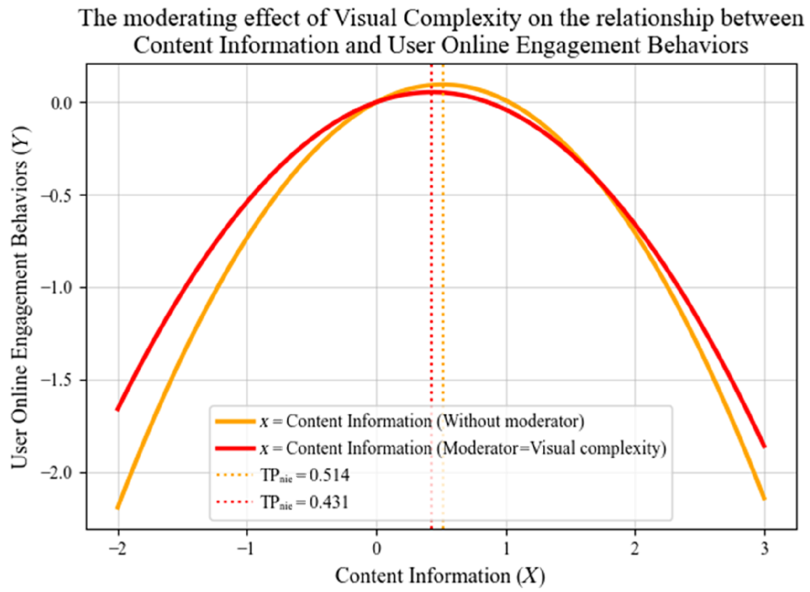

Figure 7. The moderating effect of visual complex on Content Information and User online engagement behavior. Source: Authors own work

In Model 5 (see Table 6), we also tested the interaction effect between text sentiment and visual sentiment on UOEB using Equation (7).

$$\begin{aligned} UOEB = {} & \beta_{14} + \beta_{15} Text\ sentiment + \beta_{16} Visual\ sentiment \\ & + \beta_{17}(Text\ sentiment \times Visual\ sentiment) + \delta_7 Text\ length + \delta_8 Chat\ number \\ & + \sum_{i=1}^{24} \gamma_i Type_i + \varepsilon_5 \end{aligned} \tag{7}$$

The independent effect of visual sentiment on UOEB was significantly negative ($\beta_{16} = -1.008$, $p < 0.05$). However, the interaction term between text sentiment and visual sentiment was significant and positive ($\beta_{17} = -0.035$, $p < 0.05$). This indicates that the combination of images with lower positive emotional levels and text with positive sentiment enhances user online engagement behavior, supporting H4.

5. Conclusion

5.1 Discussion

This study integrates DCT and the ELM to examine how text information, visual complexity, text sentiment, and visual sentiment collectively shape UOEB. The findings yield several novel insights. First, text information volume exhibits a nonlinear, inverted U-shaped relationship with engagement, moderate levels optimize cognitive processing and enhance engagement, while both insufficient and excessive information reduce user responsiveness. This advances prior research by identifying a specific informational threshold that maximizes engagement, offering clear guidance for multimodal content design. Second, visual complexity significantly affects cognitive fluency. Simpler images facilitate information integration and engagement, whereas high complexity imposes cognitive strain and disrupts user attention (Schneider *et al.*, 2022). This highlights the value of minimalist visual design in enhancing user experience within multimodal environments. Third, contrary to conventional assumptions, text sentiment alone does not significantly influence engagement. In AI-mediated interactions, users prioritize instrumental value over emotional resonance, consistent with the dominance of the central processing route in ELM. Emotional cues appear to operate more effectively through the peripheral route, with rational heuristics outweighing affective ones in shaping user responses to AI agents. Most notably, the study uncovers a cross-modal interaction between visual and textual emotion: when visual emotional intensity is low, textual emotion becomes more immersive and persuasive; conversely, high visual emotionality suppresses engagement (Mishra *et al.*, 2024). This supports emerging evidence that excessive emotional stimuli can impair cognitive processing. Drawing on this insight, we propose the concept of an "emotional vacuum"—a strategically restrained affective environment that enhances the impact of textual emotion—as a design principle for AI interfaces. Together, these findings deepen theoretical understanding and offer actionable strategies for affective computing and interactive marketing.

5.2 Theoretical contributions

This study introduces a dual-channel framework that unites DCT and the ELM. Moving beyond the ELM's usual focus on text, we link visual and semantic encoding to trace how informational cues (central route) and emotional cues (peripheral route) jointly shape engagement with AI agents in multimodal settings. This synthesis enlarges ELM's explanatory reach and clarifies the cognitive-affective processes driving multimedia persuasion, showing that engagement follows an inverted-U pattern rather than the linear rise often assumed (Al-Youzbaky *et al.*, 2022). We also uncover a hierarchy in emotional transmission. Users still rely mainly on textual information when selecting an AI agent, yet low-arousal images—such as neutral faces—create an "emotional vacuum" that redirects attention to the text-image blend. Text sentiment alone does not boost engagement, but image sentiment moderates its effect: pairing subdued visuals with positive text markedly lifts engagement, demonstrating that restrained imagery balances cognition and affect. Methodologically, we move past single-modal analyses — typically LDA topic modeling or manual coding (Huang *et al.*, 2022) by combining STM, VADER sentiment scores, image-entropy metrics, and a ResNet-50 classifier. This multimodal pipeline isolates textual and visual complexity and sentiment, then quantifies their separate and joint effects on engagement, offering a fuller picture of decision-making in AI-mediated environments (Wang, 2023).

5.3 Managerial implications

Based on the findings of this study, enterprises developing AI agents for virtual-character systems should focus first on controlling textual information volume and calibrating the emotional tone of accompanying images. Engagement peaks at a moderate text load; once content becomes dense, users are distracted and participation falls. Streamline copy to the essentials—especially in interactive marketing and automated customer service—so that users can absorb key points without cognitive strain. Emotional cues matter less than informational utility. When choosing an AI agent, users care chiefly about whether the text helps them solve a problem. Heavy emotional appeals can feel gratuitous, so clarity should trump sentiment. Likewise, elaborate imagery offers little benefit and may pull attention away from the task; restrained visuals that match the text's valence foster coherence and mild emotional resonance. Because AI systems now permeate high-stress environments, designers must weigh psychological impact. Information overload or intense emotional stimulation can aggravate mental fatigue. A balanced presentation that respects users' cognitive limits is therefore critical. Finally, transparency builds trust: companies should disclose how agents are designed and what they can do, aligning product development with emerging ethical and regulatory standards.

5.4 Limitations

This study has limitations in three key areas, data generalizability, user behavior conversion, and complexity assessment dimensions, which require further exploration in future research. First, the homogeneity of data sources may limit the cross-context applicability of the findings. AI interactions in entertainment-oriented platforms differ from those in more serious contexts, which may affect the generalizability of conclusions. Future research could combine experimental designs, field studies, and secondary data, as high-quality research often involves diverse data sources (Wang, 2025a), in order to offer new insights through telling an engaging and coherent story that deepens understanding (Wang, 2025b). Second, the study does not address whether engagement behaviors translate into commercial value (Aljarah *et al.*, 2024). Future work should investigate conversion rates (e.g. clicks to purchases) and their impact on profitability to inform marketing strategies. Third, the binary sentiment classification method used here oversimplifies emotional dimensions, missing complexities such as intensity and irony. Future studies could use deep learning techniques to refine sentiment analysis and examine how user heterogeneity moderates emotional impacts on engagement.

Supplementary material

The supplementary material for this article can be found online.

References

Abdi, H. and Williams, L.J. (2010), "Principal component analysis", *Wiley Interdisciplinary Reviews: Computational Statistics*, Vol. 2 No. 4, pp. 433-459.

Airoldi, E.M. and Bischof, J.M. (2016), "Improving and evaluating topic models and other models of text", *Journal of the American Statistical Association*, Vol. 111 No. 516, pp. 1381-1403, doi: 10.1080/01621459.2015.1051182.

Al-Youzbaky, B.A., Hanna, R.D. and Najeeb, S.H. (2022), "The effect of information overload, and social media fatigue on online consumers purchasing decisions: the mediating role of technostress and information anxiety", *Journal of System and Management Sciences*, Vol. 12 No. 2, pp. 195-220.

Aljarah, A., Sawaftah, D., Ibrahim, B. and Lahuerta-Otero, E. (2024), "The differential impact of user- and firm-generated content on online brand advocacy: customer engagement and brand familiarity matter", *European Journal of Innovation Management*, Vol. 27 No. 4, pp. 1160-1181, doi: 10.1108/ejim-05-2022-0259.

An, J. and Weber, I. (2021), "#greysanatomy vs. #yankees: demographics and hashtag use on Twitter", *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 10 No. 1, pp. 523-526, doi: 10.1609/icwsm.v10i1.14767.

Arapakis, I., Lalmas, M., Cambazoglu, B.B., Marcos, M.C. and Jose, J.M. (2014), "User engagement in online N ews: under the scope of sentiment, interest, affect, and gaze", *Journal of the Association for Information Science and Technology*, Vol. 65 No. 10, pp. 1988-2005, doi: 10.1002/asi.23096.

Araujo, T. (2018), "Living up to the chatbot hype: the influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions", *Computers in Human Behavior*, Vol. 85, pp. 183-189, doi: 10.1016/j.chb.2018.03.051.

Bharadwaj, N., Ballings, M., Naik, P.A., Moore, M. and Arat, M.M. (2022), "A new livestream retail analytics framework to assess the sales impact of emotional displays", *Journal of Marketing*, Vol. 86 No. 1, pp. 27-47, doi: 10.1177/00222429211013042.

Bigne, E., Simonetti, A., Ruiz, C. and Kakaria, S. (2021), "How online advertising competes with user-generated content in TripAdvisor. A neuroscientific approach", *Journal of Business Research*, Vol. 123, pp. 279-288, doi: 10.1016/j.jbusres.2020.10.010.

Biliciler, G., Raghunathan, R. and Ward, A.F. (2022), "Consumers as naive physicists: how visual entropy cues shift temporal focus and influence product evaluations", *Journal of Consumer Research*, Vol. 48 No. 6, pp. 1010-1031, doi: 10.1093/jcr/ucab042.

Blasco-Arcas, L., Hernandez-Ortega, B.I. and Jimenez-Martinez, J. (2016), “Engagement platforms: the role of emotions in fostering customer engagement and brand image in interactive media”, *Journal of Service Theory and Practice*, Vol. 26 No. 5, pp. 559-589.

Bozkurt, S., Gligor, D., Hollebeek, L.D. and Sumlin, C. (2025), “Understanding the effects of firms’ unresponsiveness on social media toward customer feedback on customers’ engagement: the impact of ethnicity”, *Journal of Research in Interactive Marketing*, Vol. 19 No. 1, pp. 59-75, doi: 10.1108/jrim-09-2023-0317.

Brodie, R.J., Hollebeek, L.D., Jurić, B. and Ilić, A. (2011), “Customer engagement: conceptual domain, fundamental propositions, and implications for research”, *Journal of Service Research*, Vol. 14 No. 3, pp. 252-271.

Bruce, N.D. and Tsotsos, J.K. (2009), “Saliency, attention, and visual search: an information theoretic approach”, *Journal of Vision*, Vol. 9 No. 3, p. 5, doi: 10.1167/9.3.5.

Chandra, S., Shirish, A. and Srivastava, S.C. (2022), “To be or not to be. . . human? Theorizing the role of human-like competencies in conversational artificial intelligence agents”, *Journal of Management Information Systems*, Vol. 39 No. 4, pp. 969-1005, doi: 10.1080/07421222.2022.2127441.

Chatterjee, P. (2012), “The role of varying information quantity in ads on immediate and enduring cross-media synergies”, *Journal of Marketing Communications*, Vol. 18 No. 3, pp. 217-240, doi: 10.1080/13527266.2011.567458.

Cheung, C. M.-Y., Sia, C.-L. and Kuan, K.K. (2012), “Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective”, *Journal of the Association for Information Systems*, Vol. 13 No. 8, pp. 2-635, doi: 10.17705/1jais.00305.

Dong, X., Veloutsou, C. and Morgan-Thomas, A. (2024), “Negative online brand engagement: conceptualisation, scale development and validation”, *Journal of Research in Interactive Marketing*, Vol. 18 No. 5, pp. 787-814, doi: 10.1108/jrim-09-2023-0303.

Eitel, A. and Scheiter, K. (2015), “Picture or text first? Explaining sequence effects when learning with pictures and text”, *Educational Psychology Review*, Vol. 27 No. 1, pp. 153-180, doi: 10.1007/s10648-014-9264-4.

Flores, W., Chen, J.-C.V. and Ross, W.H. (2014), “The effect of variations in banner ad, type of product, website context, and language of advertising on Internet users’ attitudes”, *Computers in Human Behavior*, Vol. 31, pp. 37-47, doi: 10.1016/j.chb.2013.10.006.

Gao, L., Li, G., Tsai, F., Gao, C., Zhu, M. and Qu, X. (2023), “The impact of artificial intelligence stimuli on customer engagement and value co-creation: the moderating role of customer ability readiness”, *Journal of Research in Interactive Marketing*, Vol. 17 No. 2, pp. 317-333, doi: 10.1108/jrim-10-2021-0260.

Gizzi, E., Nair, L., Chernova, S. and Sinapov, J. (2022), “Creative problem solving in artificially intelligent agents: a survey and framework”, *Journal of Artificial Intelligence Research*, Vol. 75, pp. 857-911, doi: 10.1613/jair.1.13864.

Gross, D.M. (2010), “Defending the humanities with Charles Darwin’s the expression of the emotions in man and animals (1872)”, *Critical Inquiry*, Vol. 37 No. 1, pp. 34-59, doi: 10.1086/656468.

Haans, R.F., Pieters, C. and He, Z.L. (2016), “Thinking about U: theorizing and testing U-and inverted U-shaped relationships in strategy research”, *Strategic Management Journal*, Vol. 37 No. 7, pp. 1177-1195, doi: 10.1002/smj.2399.

He, K., Zhang, X., Ren, S. and Sun, J. (2016), “Deep residual learning for image recognition”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778, doi: 10.1109/cvpr.2016.90.

Henlein, A., Bauer, A., Bhattacharjee, R., Ćwiek, A., Gregori, A., Kügler, F., Lemanski, J., Lücking, A., Mehler, A., Prieto, P., Sánchez-Ramón, P.G., Schepens, J., Schulte-Rüther, M., Schweinberger, S.R. and von Eiff, C.I. (2024), “An outlook for AI innovation in multimodal communication research”, *International Conference on Human-Computer Interaction*, pp. 182-234, doi: 10.1007/978-3-031-61066-0_13.

Hollebeek, L.D., Glynn, M.S. and Brodie, R.J. (2014), "Consumer brand engagement in social media: conceptualization, scale development and validation", *Journal of Interactive Marketing*, Vol. 28 No. 2, pp. 149-165, doi: 10.1016/j.intmar.2013.12.002.

Hollebeek, L.D. and Macky, K. (2019), "Digital content marketing's role in fostering consumer engagement, trust, and value: framework, fundamental propositions, and implications", *Journal of Interactive Marketing*, Vol. 45 No. 1, pp. 27-41, doi: 10.1016/j.intmar.2018.07.003.

Huang, G.I., Wong, I.A. and Law, R. (2021), "Dilemma of hotel reviews: the role of information processing and validation through metacognition", *Journal of Travel Research*, Vol. 60 No. 6, pp. 1301-1317, doi: 10.1177/0047287520938863.

Huang, X., Kader, M.S. and Kim, S. (2022), "Let's play with emojis! How to make emojis more effective in social media advertising using promocodes and temporal orientation", *Journal of Research in Interactive Marketing*, Vol. 16 No. 4, pp. 665-682, doi: 10.1108/jrim-04-2021-0113.

Huo, H., Wang, C., Han, C., Yang, M. and Shang, W.-L. (2024), "Risk disclosure and entrepreneurial resource acquisition in crowdfunding digital platforms: evidence from digital technology ventures", *Information Processing and Management*, Vol. 61 No. 3, 103655, doi: 10.1016/j.ipm.2024.103655.

Hutto, C. and Gilbert, E. (2014), "Vader: a parsimonious rule-based model for sentiment analysis of social media text", *Proceedings of the International Aaai Conference on Web and Social Media*, Vol. 8 No. 1, pp. 216-225, doi: 10.1609/icwsm.v8i1.14550.

Kang, H. and Lou, C. (2022), "AI agency vs. human agency: understanding human–AI interactions on TikTok and their implications for user engagement", *Journal of Computer-Mediated Communication*, Vol. 27 No. 5, zmac014, doi: 10.1093/jcmc/zmac014.

Kim, M. (2019), "Digital product presentation, information processing, need for cognition and behavioral intent in digital commerce", *Journal of Retailing and Consumer Services*, Vol. 50, pp. 362-370, doi: 10.1016/j.jretconser.2018.07.011.

Kim, J., Giroux, M. and Lee, J.C. (2021), "When do you trust AI? The effect of number presentation detail on consumer trust and acceptance of AI recommendations", *Psychology and Marketing*, Vol. 38 No. 7, pp. 1140-1155, doi: 10.1002/mar.21498.

Kitchen, P., Kerr, G.E., Schultz, D., McColl, R. and Pals, H. (2014), "The elaboration likelihood model: review, critique and research agenda", *European Journal of Marketing*, Vol. 48 Nos 11/12, pp. 2033-2050, doi: 10.1108/ejm-12-2011-0776.

Lee, J. and Hong, I.B. (2016), "Predicting positive user responses to social media advertising: the roles of emotional appeal, informativeness, and creativity", *International Journal of Information Management*, Vol. 36 No. 3, pp. 360-373, doi: 10.1016/j.ijinfomgt.2016.01.001.

Li, B. and Lima, D. (2021), "Facial expression recognition via ResNet-50", *International Journal of Cognitive Computing in Engineering*, Vol. 2, pp. 57-64, doi: 10.1016/j.ijcce.2021.02.002.

Li, Y. and Xie, Y. (2020), "Is a picture worth a thousand words? An empirical study of image content and social media engagement", *Journal of Marketing Research*, Vol. 57 No. 1, pp. 1-19, doi: 10.1177/0022243719881113.

Mir, I.A. and Salo, J. (2025), "Mapping content-driven engagement and attitudinal spillover effect of influencer marketing", *Journal of Research in Interactive Marketing*, Vol. 19 No. 3, pp. 408-425, doi: 10.1108/jrim-10-2023-0349.

Mishra, S., Moharana, T.R. and Chatterjee, R. (2024), "Exploring the role of self-conscious emotions between consumer minimalism and rental behavior", *Marketing Intelligence and Planning*, Vol. 42 No. 2, pp. 262-283, doi: 10.1108/mip-07-2023-0322.

Moreno-Munoz, A., Bellido-Outeirino, F., Siano, P. and Gomez-Nieto, M. (2016), "Mobile social media for smart grids customer engagement: emerging trends and challenges", *Renewable and Sustainable Energy Reviews*, Vol. 53, pp. 1611-1616, doi: 10.1016/j.rser.2015.09.077.

Morosan, C. and Dursun-Cengizci, A. (2024), "Letting AI make decisions for me: an empirical examination of hotel guests' acceptance of technology agency", *International Journal of Contemporary Hospitality Management*, Vol. 36 No. 3, pp. 946-974, doi: 10.1108/ijchm-08-2022-1041.

Paivio, A. (1991), “Dual coding theory: retrospect and current status”, *Canadian Journal of Psychology/Revue Canadienne De Psychologie*, Vol. 45 No. 3, pp. 255-287, doi: 10.1037/h0084295.

Pan, W., Liu, D., Meng, J. and Liu, H. (2024), “Human–AI communication in initial encounters: how AI agency affects trust, liking, and chat quality evaluation”, *New Media and Society*, 14614448241259149, doi: 10.1177/14614448241259149.

Petty, R.E. and Cacioppo, J.T. (1984), “Source factors and the elaboration likelihood model of persuasion”, *Advances in Consumer Research*, Vol. 11 No. 1, pp. 668-672.

Pittman, M. and Haley, E. (2023), “Cognitive load and social media advertising”, *Journal of Interactive Advertising*, Vol. 23 No. 1, pp. 33-54, doi: 10.1080/15252019.2022.2144780.

Riaz, N. (2019), “The effect of multimodal metaphors on social media platforms and impact on consumer purchase decisions”, University of West London.

Rigau, J., Feixas, M. and Sbert, M. (2005), “An information-theoretic framework for image complexity”, *Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*,

Roberts, M.E., Stewart, B.M. and Tingley, D. (2019), “Stm: an R package for structural topic models”, *Journal of Statistical Software*, Vol. 91 No. 2, pp. 1-40, doi: 10.18637/jss.v091.i02.

Ryu, S. (2024), “From pixels to engagement: examining the impact of image resolution in cause-related marketing on Instagram”, *Journal of Research in Interactive Marketing*, Vol. 18 No. 4, pp. 709-730, doi: 10.1108/jrim-08-2023-0262.

Santos, S., Gonçalves, H.M. and Teles, M. (2023), “Social media engagement and real-time marketing: using net-effects and set-theoretic approaches to understand audience and content-related effects”, *Psychology and Marketing*, Vol. 40 No. 3, pp. 497-515, doi: 10.1002/mar.21756.

Scheiter, K., Schüler, A., Gerjets, P., Huk, T. and Hesse, F.W. (2014), “Extending multimedia research: how do prerequisite knowledge and reading comprehension affect learning from text and pictures”, *Computers in Human Behavior*, Vol. 31, pp. 73-84, doi: 10.1016/j.chb.2013.09.022.

Schneider, S., Beege, M., Nebel, S., Schnaubert, L. and Rey, G.D. (2022), “The cognitive-affective-social theory of learning in digital environments (CASTLE)”, *Educational Psychology Review*, Vol. 34 No. 1, pp. 1-38, doi: 10.1007/s10648-021-09626-5.

Schreiner, M., Fischer, T. and Riedl, R. (2021), “Impact of content characteristics and emotion on behavioral engagement in social media: literature review and research agenda”, *Electronic Commerce Research*, Vol. 21 No. 2, pp. 329-345, doi: 10.1007/s10660-019-09353-8.

Selvi, A., Mounika, V., Rubika, V. and Uvadharanee, B. (2024), “Collegebot: virtual assistant system for enquiry using natural language processing”, *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, pp. 1407-1414, doi: 10.1109/idciot59759.2024.10467608.

Sun, C., Ye, C., Li, C. and Liu, Y. (2023), “Virtual ideality vs. virtual authenticity: exploring the role of social signals in interactive marketing”, *Journal of Research in Interactive Marketing*, Vol. 18 No. 3, pp. 430-445, doi: 10.1108/jrim-01-2023-0011.

Sussman, K.L., Bright, L.F. and Wilcox, G.B. (2023), “Digital consumer engagement: examining the impact of native advertising exposure in a social network”, *Journal of Research in Interactive Marketing*, Vol. 17 No. 4, pp. 544-561, doi: 10.1108/jrim-10-2021-0265.

Ter Stal, S., Kramer, L.L., Tabak, M., op den Akker, H. and Hermens, H. (2020), “Design features of embodied conversational agents in eHealth: a literature review”, *International Journal of Human-Computer Studies*, Vol. 138, 102409, doi: 10.1016/j.ijhcs.2020.102409.

Valenzuela-Gálvez, E.S., Garrido-Morgado, A. and González-Benito, Ó. (2022), “Boost your email marketing campaign! Emojis as visual stimuli to influence customer engagement”, *Journal of Research in Interactive Marketing*, Vol. 17 No. 3, pp. 337-352, doi: 10.1108/jrim-02-2021-0033.

van der Waal, J. (2024), “Comparing corporate Sdg reporting: multi-year, multi-region analysis via topic modeling”, Multi-region Analysis via Topic Modeling (9 April 2024).

Van Kleef, G.A. (2009), "How emotions regulate social life: the emotions as social information (EASI) model", *Current Directions in Psychological Science*, Vol. 18 No. 3, pp. 184-188, doi: 10.1111/j.1467-8721.2009.01633.x.

Wang, C.L. (2023), "Interactive marketing is the new normal", in *The Palgrave Handbook of Interactive Marketing*, Springer, pp. 1-12.

Wang, C.L. (2024), "Editorial–what is an interactive marketing perspective and what are emerging research areas?", *Journal of Research in Interactive Marketing*, Vol. 18 No. 2, pp. 161-165, doi: 10.1108/jrim-03-2024-371.

Wang, C. (2025a), "Basic but frequently overlooked issues in manuscript submissions: tips from an editor's perspective", *Journal of Applied Business and Behavioral Sciences*, Vol. 1 No. 2, pp. 139-143, doi: 10.63522/jabbs.102007.

Wang, C.L. (2025b), "Demonstrating contributions through storytelling", *Journal of Research in Interactive Marketing*, Vol. 19 No. 1, pp. 1-4, doi: 10.1108/jrim-01-2025-455.

Wang, Z., Mao, H., Li, Y.J. and Liu, F. (2017), "Smile big or not? Effects of smile intensity on perceptions of warmth and competence", *Journal of Consumer Research*, Vol. 43 No. 5, pp. 787-805.

Wang, Q., Lau, R.Y. and Yang, K. (2020), "Does the interplay between the personality traits of CEOs and CFOs influence corporate mergers and acquisitions intensity? An econometric analysis with machine learning-based constructs", *Decision Support Systems*, Vol. 139, 113424, doi: 10.1016/j.dss.2020.113424.

Wei, K.S. (2022), "The impact of barrage system fluctuation on user interaction in digital video platforms: a perspective from signaling theory and social impact theory", *Journal of Research in Interactive Marketing*, Vol. 17 No. 4, pp. 602-619, doi: 10.1108/jrim-06-2022-0160.

Yankouskaya, A., Liebherr, M. and Ali, R. (2024), "ChatGPT addiction: from support to dependence in AI large language models", SSRN 4972612.

Yim, A., Cui, A.P. and Walsh, M. (2024), "The role of cuteness on consumer attachment to artificial intelligence agents", *Journal of Research in Interactive Marketing*, Vol. 18 No. 1, pp. 127-141, doi: 10.1108/jrim-02-2023-0046.

Zhang, Z. and Zhang, L. (2025), "Most significant impact on consumer engagement: an analytical framework for the multimodal content of short video advertisements", *Journal of Theoretical and Applied Electronic Commerce Research*, Vol. 20 No. 2, p. 54, doi: 10.3390/jtaer20020054.

Zhang, X., Zhao, Z. and Wang, K. (2024), "The effects of live comments and advertisements on social media engagement: application to short-form online video", *Journal of Research in Interactive Marketing*, Vol. 18 No. 3, pp. 485-505, doi: 10.1108/jrim-02-2023-0069.

Zhou, L. and Wang, V.L. (2025), "'What's in a name?': the effect of AI agent naming on psychological ownership and responsible behaviors in the shared economy", *Journal of Applied Business and Behavioral Sciences*, Vol. 1 No. 2, pp. 145-163, doi: 10.63522/jabbs.102008.

Corresponding author
Shizhen Bai can be contacted at: baishzh1962@126.com