

Beyond Visual Realism -Understanding the Role of Verbal Cues in Generating Consumer Affinity vs. Fear toward Virtual Influencers in Social Network Advertising-

Jeeweon Wee · Namhee Yoon^{*†} · Yoon-Jung Lee^{**}

Dept. of Human Ecology & Interdisciplinary Program in Sustainable Living System,
Korea University Graduate School

^{*}Human Ecology Research Center, Korea University

^{**}Dept. of Home Economics Education, Korea University

Received July 24, 2024; Revised September 13, 2024; Accepted September 23, 2024

Abstract

Verbal cues of virtual influencers (VIs) can enhance VI anthropomorphism and persuasion effectiveness, which can increase user interactions via positive emotional arousal. This study explored the effects of verbal cues in VI social network advertising (SNA) on consumer emotions and engagement. This study employed a mixed-method approach, utilizing text mining (word frequency analysis and latent Dirichlet allocation (LDA) topic modeling) and quantitative regression analysis, to analyze user comments from YouTube videos featuring the VI, Lil Miquela. The findings revealed that consumers exhibit various emotions toward VIs, with fear being the most prominent, followed by amazement and affinity. Regression analyses further indicated the role of VI's verbal cues (e.g., voiceover) in enhancing consumers' emotional engagement with the VI, thereby improving VI SNA performance (e.g., view counts). When VIs use voiceover, positive emotions, including amazement and affinity, increase, leading to higher view counts. Conversely, the absence of verbal cues amplifies negative emotions such as fear, reducing view counts and the overall SNA effectiveness. This study underscores the importance of incorporating human-like verbal cues in VIs to foster consumer affinity and mitigate fear about the VIs. Additionally, practical implications for marketers include the strategies to leverage verbal elements to design more relatable and engaging VI SNA content.

Key words: Virtual influencer, Emotion, Verbal cue, Video content, Social network advertising

I. Introduction

Virtual influencers (VIs) are computer-generated, fictional characters designed to mimic human influencers using artificial intelligence and graphics technologies (Martin et al., 2021). Similar to human influencers, VIs also focus on producing positive content that promotes interest in brands' products and services, thus enhancing marketing outcomes such as increased

audience reach and consumer engagement. Major F&B (e.g., KFC) and luxury fashion brands (e.g., Calvin Klein, Prada, and Chanel) have already adopted VIs to promote their brands, mainly in social network advertising (SNA). For example, Lil Miquela, an Instagram influencer with over three million followers, was nominated 1 of the 25 most influential people on the Internet by *Time* in 2018; she was also invited to the Prada FW18/19 fashion show (Robinson, 2020).

From an interactive advertising perspective, VI SNA is a novel and effective method for brands to

[†]Corresponding author

E-mail: v7nami@korea.ac.kr

communicate with consumers. It is imperative for marketers, therefore, to profile a VI as a truly human-like social character. A recent marketing attempt includes implementing voiceover on VIs in video content. According to the Computers-As-Social-Actors paradigm (Reeves & Nass, 1996), anthropomorphism is fundamental to induce individuals to perceive non-human technology as real humans, thus generating positive effects, including perceived attractiveness and positive attitudes (Ahn et al., 2022), credibility (Ham et al., 2023), social presence (Pelau et al., 2021), and engagement intents (Arsenyan & Mirowska, 2021; Miao et al., 2022). This stream of research has enhanced the understanding of consumer behavior in VI marketing by expounding that anthropomorphism in an artificial intelligence device is necessary for consumer acceptance and positive psychological responses (e.g., de Visser et al., 2016; Forlizzi et al., 2007; Nowak & Biocca, 2003). Despite this contribution, the extant research has focused heavily on the physical appearance of VIs, neglecting the effect of voice as a substantial component of anthropomorphism. Research on paralinguistic cues, however, has clearly indicated that voice is an important anthropomorphic and social cue, which provides more opportunities to illustrate human traits, states, and feelings, thus fostering VI-human communication (Van Zant & Berger, 2020). For example, Finkel and Kramer (2022) found that virtual sales agents trigger significantly more consumer interactions when they are engaged in product-related discourse. Similarly, Ding et al. (2014) highlighted the positive effect of laughter in inducing positive consumer attitude toward animated human creatures. Thus, verbal cues are considered a crucial factor fostering the perceived human-likeness of VIs—and thus communication intent—by enhancing the persuasiveness of a VI's message. Therefore, VI agents are characterized by social or human-like cues (e.g., voice) that may make them more influential, which makes the relationship between VIs' voice and consumer responses crucial.

All in all, despite the popularity and potential contribution of verbal cues in enhanced consumer re-

sponses in VI SNA, relatively little attention given to the influence of verbal cues such as emotional tone or vocal pitch on VI SNA performance (e.g., Ding et al., 2014; Feine et al., 2019; Louwerse et al., 2005). This is surprising given that VIs might be perceived as more human-like, persuasive, and more relational with voiceover compared to those without as they verbally deliver product or service recommendations in a natural and playful manner (Moriuchi, 2019).

Thus, to fill this gap, this study aims to investigate consumer responses to VI SNA video content and, specifically, how emotional responses can vary depending on the presence of verbal cues (e.g., voice). In particular, we investigate the effect of different emotional responses on VI SNA performance (e.g., view counts) and how VIs' verbal cues can influence this relationship. Despite the abundance of existing studies on VI marketing (Ahn et al., 2022; Arsenyan & Mirowska, 2021; Choudhry et al., 2022; Ham et al., 2023; Kim et al., 2024; Miao et al., 2022; Yu et al., 2024), none has examined consumer phenomenon in the context of video-based VI SNA. Thus, from an exploratory perspective, we investigate the emotions consumers exhibit toward VIs in the context of VI SNA video content (e.g., YouTube) using text-mining analysis. By conducting natural language processing on YouTube user comments, we aim to directly capture state-of-the-art consumer emotions toward VIs in an organic setting rather than relying on consumers' self-report data or predetermined theories. From a practical perspective, we aim to uncover which emotional response is positively associated with advertising performance and how the implementation of verbal cues can uplift this performance. This study is expected to provide marketing insights regarding how to utilize verbal cues in designing VIs as a truly relatable and human-like social character to foster affinity with consumers. Furthermore, from a theoretical standpoint, this study can add to the CASA (Computers-As-Social-Actors) literature by expounding the effect of verbal cue in promoting positive VI-human interactions and thus, successful VI SNA video performance.

II. Theoretical Background

1. Virtual influencers (VIs) and video content in social network advertising (SNA)

Virtual entities such as VIs are becoming increasingly useful for marketers in promoting products and services in today's retail setting. Miao et al. (2022) proposed a theory of avatar marketing, defining VIs as "digital entities with anthropomorphic (i.e., form realism), controlled by a human or software, that are able to interact (i.e., behavioral realism)" (p. 5). This definition highlights two key components of avatars: form realism and behavioral realism. Form realism refers to how much a virtual avatar resembles a human (Nowak & Fox, 2018), encompassing design elements like anthropomorphic appearance (e.g., cartoonlike character such as "Noonoouri" vs. fully human-like character such as "Lil Miquela"), spatial dimensions (e.g., 2D vs. 3D), and movements (e.g., static vs. dynamic). On the other hand, behavioral realism focuses on how authentically a virtual avatar mimics human behavior (Guadagno et al., 2007), including elements such as communication modalities (e.g., verbal communication via speech vs. non-verbal communication through text, gestures, or facial expressions vs. a combination of both), response type (e.g., scripted vs. natural), and the presence of social content (e.g., engaging in discussion about social and personal matters vs. limited to task-oriented communication). Although these components jointly determine the extent to which a virtual influencer can resemble and behave like a human, influencing their effectiveness in fostering engagement and relatability in marketing contexts, behavioral realism elements are particularly important for featuring a VI as a truly authentic, autonomous, and social human character; it can have its own personality, be able to express emotions, and interact with consumers through social networking sites based on behavioral realism elements (Lou et al., 2023). This makes VIs notably different from other task-based digital marketing agents who tend to be limited to responding to typical consumer requests (e.g., online sales/recommendation

agents or chatbots) than engaging in two-way communication.

Previous studies have demonstrated that media post modality is an important consideration for designing VIs' behavioral realism. Specifically, video content (vs. static images) enriches the human elements of VIs more effectively and thus generates greater emotional arousal from users. For the VI's self-related information (e.g., posts containing VIs' daily routine, personal thoughts, and opinions), in particular, perceived information richness is amplified when shared in video posts compared to images (Li et al., 2019). DataBridge (2022) also showed that VI video posts gained higher engagement rates (e.g., liking, commenting, or sharing counts) than the static image posts of VIs. Supported by this academic validation, marketers are increasingly utilizing various forms of video content in the promotional SNA campaigns of VIs. Furthermore, marketing effectiveness of social media posts relies heavily on consumer engagement and interactions from the audience (Dwivedi et al., 2023); consumer engagements such as viewing, commenting, and sharing play a crucial role in fostering a sense of connection that enhances the credibility of posts and cultivates loyalty toward the brands endorsed by VIs (Chidiac & Bowden, 2023). When followers engage in liking, commenting, or sharing, it increases the spread of the content to a wider audience, which thus attracts more followers and potential customers.

2. Emotional responses to VI SNA video content

Virtual humans in social media posts can evoke various emotional reactions among users. This may be natural in that when such entities are more human-like than other machines, we may be able to relate to them and even feel an emotional connection. An affective response is primarily transient emotive feedback that results from an individual's exposure to an emotionally-arousing stimulus (Wilson et al., 2001). Users can experience a wide range of positive feelings as they encounter virtual humans, including fun, amusement,

or wonder (Koles et al., 2024), curiosity or novelty (Dabiran et al., 2023; Franke et al., 2023), and affinity (Dabiran et al., 2023). Among various positive emotions toward VIs, affinity may most comprehensively capture the rest of the various positive feelings toward them (Bartneck et al., 2007; Bartneck et al., 2009) by encompassing likeability, familiarity, pleasantness, and empathetic responses (Wang et al., 2015). Additionally, recent studies have pointed to consumers' awe-inspired affects evoked by VI media content, particularly by a unique mixture of visual appeal, sense of mystery, and creative storytelling that sets VI content apart from that of human influencers. According to Choudhry et al. (2022), this emerging affect could be distinguished from affinity, which is a general emotional favorability to VIs due to familiarity and human-likeness, but could be captured in a separate emotional term, possibly as amazement. Choudhry et al. (2022) interviewed 30 participants who commented on VIs' Instagram posts and reported that followers find VIs funny, entertaining, and cool. The creativity, novelty, and visual attractiveness of the content are key drivers of engagement with VIs.

Substantial research has highlighted that positive emotion is a substantial factor in fostering VI SNA performance. Kim et al. (2024) showed that positive feelings of coolness toward VIs mediated the realistic VI interface designs and purchase intents of VI-endorsed products. Lim and Lee (2023) also demonstrated that emotion-arousing media content posted by VIs increases their credibility via parasocial interactions. Arsenyan and Mirowska (2023), in a similar vein, noted that viewers' sense of empathy determines VI SNA performance, fostering the degree of interactions with VIs (e.g., following intents) by making them appear more socially attractive.

As human-like entities become increasingly similar to humans in appearance, however, they may evoke repulsive reactions in humans, which is known as the uncanny valley effect. This phenomenon arises due to media users' inherent resistance to artificial entities. People tend to doubt that computer-generated entities can behave and function like real humans due to lack

of fundamental human traits such as agentive reasoning capacity or emotional experience (Longoni et al., 2019). Consequently, individuals are likely to experience uncanny feelings toward highly realistic VIs, especially when one can know about their non-human nature. The uncanny valley refers to a state of psychological discomfort or unease when encountering almost human-like, yet not fully human-like entities (Mori et al., 2012). The nuanced gap between "almost human" and truly indistinguishable human features tend to provoke feelings of eeriness, encompassing creepiness, fear, shock, disgust, and repulsion. Furthermore, an array of these negative feelings, in turn, lead to reduced users' affinity toward these entities (Ho & MacDorman, 2010; Mathur & Reichling, 2016). More recent studies have added that a linkage exists between VIs' interface realism and the arousal of such negative emotions toward VIs. For example, Kim et al. (2024) demonstrates that a feeling of eeriness increases in viewers owing to a lack of human-like behavioral attributes of VIs, which in turn leads to decreased intent to purchase the VI-endorsed product.

In the meantime, uncanniness is hardly comprised of a single emotional experience, but rather a nexus of emotional phenomena with disparate psychological and perceptual mechanisms. Nonetheless, this feeling essentially contains fearful emotion (Wang et al., 2015). MacDorman et al. (2009) contended that uncanny feeling is deeply rooted in fear rather than disgust, and ascribed the uncanny phenomenon to the mechanism of pathogen avoidance in search for human survival (MacDorman et al., 2009; MacDorman & Ishiguro, 2006). Particularly, fear is evoked with the entities' increased human-likeness because, as these entities are perceived to be genetically closer to human, their imperfections elicit stronger aversion than those found in their mechanical-looking counterparts. Ho et al. (2008) also argued that uncanny emotion is associated with fear, but from a different angle. They argued that human replicas naturally remind humans of death, thus eliciting uncanny feelings driven by fear of mortality. Indeed, many humanoid replicas—including dolls, clowns, or mannequins—resemble dead individuals

who have seemingly come back to life. VIs may also be perceived as threatening, as they evoke fears related to being replaced by an android doppelganger, losing autonomy over one's physical presence, or being stripped of the essence of a soul. These results collectively suggest that fear underlies the eeriness or creepiness that humans attribute to VIs.

3. Verbal cues in VI SNA

Communication modality is a social cue for humanifying VI characters in SNA videos (Guadagno et al., 2007). Specifically, a voiceover that provides information, context, or commentary is a common form of communication modality used in VI SNA. Communication cues drive perceptions of virtual assistants and their influence on social, relational, and persuasive outcomes (e.g., Cheng et al., 2021; Lee et al., 2021; Rhee & Choi, 2020). As a type of communication cue, a voiceover can implement human-likeness of virtual agents and enhance their persuasiveness.

According to the CASA paradigm, humans tend to apply social norms to their interactions with technology in a manner similar to their interactions with real human beings. This framework suggests that implementing communication cues to computerized agents can therefore, enhance social responses from human audience. This is because since the use of voice is an inherently human trait (Pinker & Bloom, 1990), a VI emulating the human is perceived as more human-like than if they were to communicate via text only. This framework is further corroborated by Schroeder and Epley (2016), who demonstrated through a series of experiments that paralinguistic features of speech, such as intonation and pace, shape perceptions of human-likeness more effectively than text-based interaction. Unlike text, voice allows for a richer and more realistic portrayal of human traits, emotional states, and feelings during communication (Van Zant & Berger, 2020).

Voice can thus influence VIs' message persuasiveness via perceptions of human characteristics (e.g., Van Zant & Berger, 2020). Voice in itself (vs. text)

might be the element that gives virtual human agents their human-like appeal (Cho et al., 2019); this may enhance their social character and, in turn, their persuasiveness. Indeed, empirical studies have revealed that voice-based assistants are more persuasive than text-based assistants because they can deliver product or service recommendations verbally in a natural manner (Moriuchi, 2019). These verbal elements must be crafted to positively influence the audience's favorable perceptions of and behaviors toward VIs, thereby enhancing the overall impact of VI marketing.

The literature has broadly assumed that realism is a desired outcome for anthropomorphic entities. VIs are thus required to have a highly sophisticated form and behavioral realism compared to other animal- or cartoon-like entities. For instance, while a police dog could be anthropomorphized as a brave warrior with no need to be involved in verbal discourse, chatbots are expected to manage more human-like dialogue features and empathetic capacity through written language and voice (Fernandes & Oliveira, 2021). This suggests that the anthropomorphism of VIs is conveyed not only through visuals and human-like appearances but also through human-like, social and interactive behaviors. According to Finkel and Kramer (2022), human-like robots triggered more social interactions in consumers when they were highly engaged in an interactive dialogue on product consumption. This implies that VIs can generate positive relational outcomes for consumers with heightened discourse engagement. Verbal engagement is, therefore, thought to be a substantial factor in fostering the authenticity of VIs as real social characters. However, it also strengthens the persuasive effect, which can, in turn, earn positive consumer responses.

Based on the above, this study poses two important research questions: RQ (1) To what extent do consumers' emotional responses to VIs vary? RQ (2) How do these emotional responses vary depending on the presence or non-presence of verbal cues? To address RQ (1), we identify consumers' emotional response to VIs using topic modeling and word frequency analysis conducted on the user comments of YouTube VI SNA

content. Next, to address RQ (2), we conduct regression analysis to examine (1) how emotions influence VI SNA performance (e.g., view counts) and (2) how the influence of emotions on VI SNA performance varies depending on the presence or absence of VI's verbal cues.

III. Methodology

1. Data-collection process

We collected a set of user comments on YouTube video content to explore consumers' emotional and cognitive responses to VI SNA video content. Users on YouTube provide perspectives, feelings, and thoughts to express and form the basis of their opinions. Owing to the volume of user comments serving as a resourceful dataset, user comments are the most frequently utilized in data analysis (Loosen et al., 2018). Therefore, analyzing user comments is beneficial for understanding their organic thoughts and emotions related to VI SNA video content.

The data were collected in May of 2023. We used the keyword "virtual influencer" with the language settings restricted to English. The initial dataset comprised 25 videos. Most of the collected videos featured Miquela as the VI. To control for the confounding influence of different VI characters, we excluded six videos that featured other characters, such as Imma and Nounoori. In support of our decision to use Miquela as a representative VIs, numerous studies and trade publications have noted the prominence of Lil Miquela in the social media marketing arena worldwide (e.g., Block & Lovegrove, 2021; Dabiran et al., 2023; Kim et al., 2024). For example, Block and Lovegrove (2021) examined Miquela as the most commercially successful VI, reporting identity intrigue, discordant storytelling, emotional release, and provocation as crucial communication strategies that foster consumer relationship-building. Miquela is a popular Brazilian-American lifestyle influencer and singer from LA, with approximately 8 million followers on social media. She is less character-like but is fairly akin to human appear-

ance in terms of the level of anthropomorphism (Kim et al., 2024). With such highly human-like appearance, she also features socially human-like features such as taking a role as a brand ambassador for the major fashion brands such as Chanel, Calvin Klein, and Prada; she also boasts her friendship with the American supermodel Bella Hadid through social media posts. The Time magazine's nomination of her as one of the most influential people on the Internet also contributes to the building of Lil Miquela's identity as a socially-capable and -interactive digital influencer (Koles et al., 2024).

2. Data information

Through the above selection process, we collected 19 videos (10 videos with voiceover and 9 videos without) ranging from commercials to self-presentation (e.g., romance, break-up stories, or personal insecurities) or music videos created to promote Miquela as a fictitious social, live character and build relationship with viewers. We collected comments on selected videos using a web crawler in Python. The crawled dataset included date, title, comments, comment likes, video likes, and view counts of each show. The view counts of the collected video shows ranged from 51,212 to 7,121,949 and likes ranged from 1,300 to 98,000. We collected 27,220 comments and analyzed them using text mining.

3. Data pre-processing

We performed natural language pre-processing to analyze user-generated comments, which required data conversion into a structured format. The text was cleaned by removing numbers, non-English characters, and meaningless punctuation. This was followed by the removal of "stop words," which are non-informative words (e.g., "a(n)," "the," "is," "of," "I," "we"). Additionally, we reduced words to their root forms to avoid counting synonyms multiple times. After text pre-processing, we conducted frequency analysis and topic modeling analysis to extract meaningful information and knowledge.

IV. Results

1. Term Frequency-Inverse Document Frequency analysis: Overall consumer response to VI SNA content

As a preliminary analysis of consumer responses to VI content, we conducted a Term Frequency-Inverse Document Frequency (TF-IDF) analysis—a frequency analysis method intended to enhance the ability to differentiate documents by weighting unique and significant terms more heavily. Based on the TF-IDF analysis, we derived the 50 most frequent words (Table 1). According to the results, the term “Robot” was most

frequently mentioned in comments; this preponderance indicates the word’s frequent association with VIs. This suggests that viewers are highly aware of the non-human identity of these influencers and that their robotic nature is central to the viewer experience. The words “Song” and “Voice” also featured prominently, which indicates that auditory elements are substantial in VI content. These terms suggest that music and vocal performances play a crucial role in how these digital personas engage with the audience; this highlights the effectiveness of strategic use of multimedia elements in creating VI SNA video content.

Next, we extracted nouns related to the VI’s appearance, such as “Face,” “Body,” “Look,” “Hair,” and

Table 1. Top 50 words on VI YouTube comments and TF-IDF results

No.	Word	TF-IDF	No.	Word	TF-IDF
1	Robot	906.521475	26	Friends	125.786849
2	Song	600.391126	27	LMAO	123.711207
3	Miquela	558.399335	28	Shane dawson	118.771619
4	Voice	375.677433	29	Cool	116.209848
5	Love	337.387488	30	AI	114.239969
6	Face	322.965337	31	Guy	114.013691
7	Girl	297.430569	32	Character	112.688700
8	Omg	270.611532	33	Fuck	110.491695
9	Lol	247.192796	34	Shit	107.692153
10	Animation	246.184207	35	Question	106.709726
11	Wtf	238.586698	36	Understand	104.871941
12	Wow	236.797184	37	Video	104.427539
13	Bop	230.241441	38	Amazing	102.093130
14	Life	219.586379	39	Age	98.794733
15	Body	210.339750	40	God	97.234847
16	Human	209.517113	41	Live	94.590078
17	Look	205.699384	42	Hair	91.475056
18	See	158.807727	43	Hope	91.384459
19	People	156.565351	44	Pretty	88.064885
20	Music	153.685589	45	Fake	85.474511
21	Real	153.189860	46	Happy	84.064154
22	Robots	144.608853	47	Hate	84.002707
23	Cute	127.225655	48	Humans	83.494274
24	Confused	125.862911	49	Filter	81.812204
25	Beautiful	125.798892	50	Eyes	81.551313

“Eyes,” which indicates that viewers focused their attention on VIs’ physical features for eliciting human interaction and recognition. For VIs, a well-designed face that can convey emotions and expressions may be crucial in creating a relatable and engaging persona. The frequent mention of these physical aspects of VIs suggests that audiences are closely observing them, possibly comparing them to human influencers or assessing their realism and appeal.

Last, emotional expressions such as “Lol,” “Wow,” “Confused,” “Cool,” “Amazing,” and “Happy” were highly dominant. The prevalence of positive terms indicates that VIs successfully foster positive emotional connections, enhancing brand affinity and viewer loyalty. The presence of humor-related terms such as “LMAO” suggests that VIs effectively incorporate entertaining elements, making the advertisement content more enjoyable and shareable.

2. Latent Dirichlet allocation (LDA) topic modeling analysis: Emotional response to VI SNA video content

We adopted the latent Dirichlet allocation (LDA) machine-learning method to address the emotional responses discussed among the comments to VI SNA video content (RQ 1). The LDA can reflect a probabilistic distribution of words and helps to extract latent topics based on this distribution (Lucini et al., 2020). In this study, we employed the Python Genism library for the LDA analysis. We manually labeled the topics ex-

tracted from the LDA results and categorized them into consumers’ cognitive or emotional appraisal. To verify the interpretation of topic labels, we reviewed the original comments containing these keywords to ensure that the labeled topics accurately reflected their original meaning.

As a result of applying the LDA model, we identified eight distinct topics, each representing major responses to VI SNA video content (Table 2). “Fear” was the most prominent (20.64%); in this topic, we identified words such as “Creepy,” “Freaking,” and “Uncanny.” Topic 2 was named “Amazement” based on words such as “Wow,” “Super,” and “Stunning.” These words reflect a positive emotional state, such as awe or amazement (18.66%). The next category focused on the setting or environment surrounding VIs. Topic 3 was labeled “Virtual environment” because words such as “Game,” “Fortnite,” and “Songs” represent the characteristics of the virtual stage where VIs perform. Topic 4 revolved around consumers’ appraisal of VIs’ appearance and was labeled “Appraisal for appearance.” This cluster comprises words such as “Gorgeous,” “Cute,” and “Pretty,” representing consumers’ positive evaluations of the VIs’ physical attractiveness. Topic 5 was identified as another cluster for emotional responses and named “Affinity” based on the words such as “Love,” “Adore,” and “Crush.” The following important category, Topic 6, was named “Anthropomorphism,” based on the words “Robot,” “Face,” and “Eyes.” This topic reflects the fact that consumers not only evaluate the physical attractiveness of VIs, as in

Table 2. Results of LDA topic modeling

Topic	Keywords	%
Fear	Weird, Don’t, Confused, How tf, How?, Creepy, Freaking, IDK, Scared, Wonder, Uncanny	20.64
Amazement	Wow, OMG, Amazing, Super, Surprise, Believe, Stunning, Wonderful, Incredible	18.66
Virtual environment	AI, Game, Fortnite, Music, Catwalk, Video, Song, Skin, Play	13.44
Appraisal for appearance	Nice, Beautiful, Cute, Look, Gorgeous, Pretty	13.11
Affinity	Love, Adore, Hug, Fan, Crush, Happy, Like	11.93
Anthropomorphism	Robot, Human, Doll, Toy, Hair, Eyes, Body	10.08
Gender	Girl, Guy, Dude, Bro, Gal, Princess	6.55
Social agenda	Anti, Earth, Feminism	5.21

Topic 4, but they also share their thoughts about VIs' visual qualifications for realism. The last two topics, 7 (6.55%) and 8 (5.21%), showed a marginal percentage compared to the other top-ranked topics. Although marginal, Gender and Social agenda have emerged as new topics. This finding, which could have been underestimated in the literature, sheds light on new research interests. Collectively, over 50% of the total discussed topics center around emotional responses (fear, amazement, and affinity), indicating the importance of emotions in consumer experience with VI SNA. This was followed by a cluster of Appearance appraisal and Anthropomorphism, which take up 23.19% combined.

The three major emotions extracted from the topic modeling analysis are fear, amazement, and affinity. The "Fear" topic, which comprises 20.64% of the analyzed comments, is dominated by emotion-laden terms that express discomfort and anxiety. Words such as "Confused" and "Creepy" suggest a strong emotional reaction to the lifelike but artificial nature of VIs. The confusion reflected in terms such as "How tf" and "How?" indicates that viewers are grappling with the unfamiliarity and strangeness of digital personas. The emotions of fear and confusion point to the uncanny valley effect, where VIs are perceived as eerily close to human yet not quite human, which leads to discomfort. This highlights the need for marketers to address and mitigate these emotional responses to make VIs more palatable to a broader audience.

Conversely, the "Amazement" topic, accounting for 18.66% of the comments, is replete with terms that convey positive surprise and wonder. Words such as "Wow," "OMG," and "Stunning" express a strong emotional reaction to the visual and technological aspects of VIs, coupled with a sense of fun and curiosity. These terms suggest that many viewers are not only impressed but also genuinely astonished by the capabilities and presentation of these digital figures. The emotion of amazement is critical in capturing audience interest and can be leveraged to enhance the novelty appeal of VIs in marketing campaigns.

As another positive emotional term, the "Affinity" topic represents 11.93% of the comments. This topic

can, however, be differentiated from "Amazement" in that it is characterized by words that reflect clearly deeper affection toward VIs. Terms such as "Love," "Adore," and "Crush" indicate that some viewers have developed strong emotional bonds with these virtual personas, akin to the relationships formed with human influencers. The use of words such as "Happy" and "Fan" suggests that these VIs bring joy and satisfaction to their followers, fostering a loyal and enthusiastic fanbase. Particularly noteworthy from frequency analysis is the cooccurrence of positive and negative emotions in user comments, which illuminates the ambivalence in consumer feelings toward digital humans. This notes the need for marketing strategies that can help mitigate VIs' artificial feature while enhancing authentic and relatable interactions. More detailed discussions continue in the Conclusions and Implications.

3. The effect of VIs' verbal cues on VI SNA video content

To examine how the emotional responses influence viewing depending on the conditions of providing or not providing VIs' verbal cues, we tested a multiple-regression model. First, to formulate independent variables, we coded comments into numerical scores for each type of emotion. Based on the results of topic-modeling analysis, we identified three emotion categories (affinity, amazement, and fear) and the relevant words representing each emotion category. We then coded the comments based on whether they contained words from these emotion categories (0 if the words are not present; 1 if the words are present), and we summed the scores by emotion category. For example, if a comment contained the word "Creepy," this comment was coded as 1 for fear. If a comment included the word "Amazing," it was coded as 1 for amazement. In case of a comment such as "Oh, what a freaking face, lol can't believe it's a 19-year-old digital girl. So creepy but truly amazing thou," the comment as a unit of analysis contained two words ("Freaking" and "Creep") related to fear and one word ("Amazing") related to

amazement. Therefore, this comment was coded into fear (2), amazement (1), and affinity (0). Coding was conducted by the first and second authors. The two coders finalized the coding until they reached a consensus. The number of agreements was divided by the number of agreements plus disagreements, yielding an intercoder reliability of .92. Following this process, the scores for each emotion category (affinity, amazement, and fear) were calculated.

A total of 5,059 comments included at least one valid emotional response. We used the scores for the three emotion categories as independent variables and content view counts as the dependent variable. We performed regression analysis using SPSS 26. <Table 3> shows means and standard deviations of each of the four variables analyzed and the correlations among those. All the variables were significantly correlated. Meanwhile, the pairwise correlations among the predictor variables (amazement, affinity, and fear) showed low correlations, indicating a potentially low multi-

collinearity issue.

According to the multiple regression analysis, amazement, affinity, and fear significantly influenced view counts (Table 4). Amazement had a positive influence on view counts ($\beta = .17, t = 9.92, p < .01$). Affinity also had a significant positive influence on view counts ($\beta = .29, t = 2.27, p < .001$). On the other hand, fear had a significant negative influence on view counts ($\beta = -.24, t = -11.95, p < .001$).

Subsequently, regression analyses were conducted for two separate conditions: 1) 10 shows (with 3,052 comments) containing VIs' verbal cues (e.g., voiceover) and 2) 9 shows (with 2,007 comments) not containing VIs' verbal cues. Voiceover takes the form of dubbing an auditory message on VIs, which is recorded by a third-party voice actor outside the scene. In the videos used in our study, most voiceovers were provided with background music. In both the conditions, text captions were not provided. The results (Table 4) showed that for the condition in which verbal cues

Table 3. Correlations among variables

Variables	Mean (SD)	1	2	3	4
View counts	2,966,801.03 (2,425,307.29)	1			
Amazement	0.16 (0.39)	.13*	1		
Affinity	0.48 (0.54)	.17**	.21***	1	
Fear	0.25 (0.45)	-.18**	-.18**	-.32***	1

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4. Results of multiple regression analysis

Condition	Independent variables	β	F-ratio	R ²
Total (N=5,059)	Amazement	.17**	77.91***	.17
	Affinity	.29***		
	Fear	-.24***		
Shows with verbal cues (N=3,052)	Amazement	.17*	31.70***	.14
	Affinity	.38***		
	Fear	-.11		
Shows without verbal cues (N=2,007)	Amazement	.02	17.83***	.13
	Affinity	.02		
	Fear	-.29**		

Dependent variable: View counts

* $p < .05$, ** $p < .01$, *** $p < .001$

were present in the video content, fear had no significant effect on view counts. However, amazement ($\beta = .17, t = 11.472, p < .05$) and affinity ($\beta = .38, t = 20.95, p < .001$) had significant positive effects on view counts. Interestingly, the effect of the affinity substantially increased under these conditions. This implies that negative emotional responses diminished, while positive emotions, such as awe or affinity, increased when VIs actively engaged in verbal discourse, thus leading to increased video content view counts. When verbal cues were not present, only fear had a significant negative effect on view counts ($\beta = -.29, t = -10.818, p < .01$). Unlike the condition with VIs' verbal cues, affinity and amazement had no significant effect on view counts. These findings imply that when VIs perform visually without dialogue expressions, they may be perceived as more fearful or strange, thus negatively impacting content viewing.

V. Conclusions and Implications

This study utilized text-mining and regression analysis to examine consumers' emotional responses to VI SNA videos using user-generated comments of YouTube content. We first identified consumer emotions to VI SNA video content in a naturalistic setting using a text-mining approach and quantitatively examined how these emotions influence content viewing, particularly in relation to whether VIs' voiceover is provided as a verbal cue. This study offers practical implications for enhancing VI SNA effectiveness.

The results of TF-IDF analysis provided a meaningful snapshot of consumers' general attention toward or interest in the VI video content. The analysis revealed the top 50 words frequently mentioned, including key terms associated with VIs' appearance and identity, such as "Robot," "Voice," "Face," "Girl," "Body," "Look," and "Hair," followed by the next key terms associated with emotional expressions, such as "Lol," "Wow," "Confused," "Cool," "Amazing," and "Happy." Overall, consumers concentrated on VI's physical features or their non-human identity, and this was accompanied by emotional evaluations.

For the further text-mining analysis, we conducted LDA topic modeling and extracted eight distinct topics, with each topic providing insight into the various dimensions of consumer responses to VI content. The most significant and overarching topic was consumers' emotional responses, which indicates that emotion plays a pivotal role in shaping consumer experience with VIs. This echoes previous research focusing on the formation of emotional perceptions in various human-technology interaction contexts (e.g., Bailenson et al., 2006; Slater & Banakou, 2021). These studies have emphasized the importance of developing consumer rapport (Gratch et al., 2007) and empathy (Bailenson, 2018; Parmar et al., 2022) in generating positive behavioral intents toward VIs. Nonetheless, the findings reveal that fear remains the most prominent emotion, consolidating the uncanny-valley effect in the context of digital influencers as well. According to the preponderant earlier work built on the uncanny-valley effect, the extreme human-likeness of digital humans provokes cognitive confusion and feelings of discomfort (Cheetham et al., 2011; Kim et al., 2024; Mathur & Reichling, 2016). Although VIs have become more visible in the recent social media marketing scene, the evocation of negative perceptions and feelings seems inevitable, thus requiring a call for VI design or communication strategies that can alleviate such negative psychological responses. As the auditory elements of VI SNA video content, such as VIs' voice or song, appeared as important terms according to the TF-IDF analysis, the use of multimedia elements that can amplify the sensory experience can be crucial. Fortunately, however, as evidenced in the previous literature (Dabiran et al., 2023; Franke et al., 2023; Kim et al., 2024), users also reported positive responses—including amazement and affinity—as they encounter virtual humans. This largely concurs with Kim et al. (2024), who showed that two countervailing perceptions of VIs, coolness and eeriness, coexist in consumers as they experience VIs. According to Warren et al. (2019), positive consumer perceptions, such as coolness, can result in successful marketing outcomes, such as consumers' willingness to pay for a

brand. In contrast, negative perceptions such as eeriness increase the likelihood of adverse behavioral intentions, especially within the context of chatbots (Pavlidou, 2021). Similar outcomes are highly likely within the context of VIs as well, such that consumers' fear toward a VI can impede positive behavioral intentions toward the VIs such as a following intent or purchase intent, while amazement and affinity can foster these behavioral intentions.

Follow-up regression analyses revealed that amazement and affinity positively influence content view counts, while fear negatively influences content view counts. These results imply the importance of crafting VI SNA content that can arouse positive emotions. This study suggests that the verbal features of VIs can specifically enhance these effects. Our findings align with those of Finkel and Kramer (2022), who asserted that favorable attitudes toward the virtual agents can be shaped by the increased level of anthropomorphism of VIs. Their study revealed that human-like robots significantly increased social interactions among consumers when they engaged in interactive dialogue about product consumption. This echoes previous studies contending that avatar-user engagement can be improved by elements other than visual ones, such as discourse involvement (e.g., Banks & Bowman, 2016). In this regard, the proper verbal featuring of VIs in SNA video content can be a driver of affinity. When VIs are only visually present, consumers may focus more intensely on their physical appearances, which intensifies fear of the hyper-realism look of artificial humans (Brenton et al., 2005). Conversely, when VIs do not engage in any dialogue, this can weaken their human-likeness owing to a lack of social character, which amplifies reluctance toward VIs. However, different findings also exist, as demonstrated by Ischen et al. (2022). According to their study, virtual assistants' verbal information did not suffice to enhance anthropomorphism. Hence, simply using voice as a human-like cue presumably was not sufficient for endowing human-like appeal to VIs and thus failed to address negative fearful responses. In summary, findings regarding the effect of verbal cues in inducing positive

responses to VIs among consumers are somewhat mixed. Hence, future studies are needed to thoroughly investigate the effect of VIs' verbal engagement in the context of video-based VI SNA.

Theoretically, this study contributes to the growing literature on VI marketing by elucidating the consumers' emotional experiences with the VIs within the video-mediated VI SNA context. First, it was conducted from an exploratory perspective. Accordingly, we used text-mining analysis to identify consumer responses directly from an unstructured, organic setting since a big-data analysis allows researchers to explore the target phenomenon as it is without reducing it based on existing theories. Future follow-up studies using empirical methods (e.g., surveys or experiments) will, however, be still valuable to reexamine the findings of the current study, and validate them within the context of video-based VI marketing.

More crucially, by building on the CASA paradigm within the context of human-technology interactions, this study directly explores how the implementation of voice impacts consumers' response to VI and/or content. While much existing research has concentrated on the human-like qualities of VIs, verbal cues have been largely overlooked. However, the current research suggests that media users respond to virtual humans differently depending on the verbal manner in which they are presented; this indicates that verbal cues can influence users' attitudes and behaviors via positive emotional arousals. This study sheds light on the mechanisms behind the emerging phenomenon in human-technology interaction, which may become more common as loneliness and the decline of friendships increase among younger generations (Pittman & Reich, 2016). When users struggle to establish social connections with real people, they may transfer the need for connectedness, or companionship to these digital entities that can instead provide emotional engagement.

Furthermore, we suggest extending the findings of this study to marketing communication theories, particularly utilizing persuasion knowledge model (Friestad & Wright, 1994). Emotions have been studied extensively in relation to persuasion and consumer atti-

tudes (Petty et al., 2003), such that emotion can influence the processing mode and the subsequent persuasion effect (Griskevicius et al., 2010). Hence, future studies in the context of virtual human interactions could incorporate cognitive load and persuasion knowledge to understand how the use of voice shapes perceived persuasiveness and its subsequent effects (e.g., brand memory or attitudes, and purchase intention) via emotional mediations. Given the growing popularity of VIs used in marketing domain, such as product recommendations, this stream of future research could illuminate the process by which consumers feel and process the messages by VIs, thus contributing to the knowledge of human-VI interactions.

Despite its contribution, this study has limitations with regard to data collection. The data-collection time is particularly crucial in big-data analysis as it directly impacts the relevance, accuracy, and effectiveness of the insights generated from the data. The data collected at a specific time point can provide insights only contextually relevant to that time point (Gandomi & Haider, 2015). Since patterns and meanings can change or evolve over time, future studies will be valuable to follow up and update the understanding of consumer behavior in this VI SNA scene. Furthermore, future studies may need to incorporate various VI characters. This study only used Lil Miquela owing to her prominence as a successful VI; future research could include various types of VIs across different age, gender, and various appearance or VI interface features.

Marketing implications are as follows. Our study confirms and extends the notion that perceived human-likeness can positively influence relational outcomes, such as generating amazement and affinity. In the context of emerging technologies, a higher degree of “human appeal” is often assumed to lead to more positive emotional outcomes (e.g., Sivaramakrishnan et al., 2007). Therefore, fashion brands should consider incorporating verbal features in VIs in the SNA video content to enhance affinity and amazement, which can subsequently lead to increased content viewing and better engagement. Indeed, VIs provide fashion brands with a novel and highly controlled way to pro-

mote their products and consumer engagement. They can be styled, posed, and directed to perfectly embody a brand’s aesthetic, possibly allowing for precision-targeted marketing campaigns that resonate with fashion-consumer demographics. Furthermore, since VIs often have large, dedicated followings on social media, which makes them powerful tools for audience engagement, fashion brands will also be able to strengthen their connection with consumers in a more interactive and relatable way by engaging with these audiences through virtual influencers.

To evoke amazement and affinity, fashion brands and content creators should feature VI characters in innovative and striking ways to foster consumers’ emotional arousal. For example, fashion brands should design video content highlighting awe-inspiring aspects of VIs using verbal techniques to enhance the social and human characteristics of VIs. Fashion VI SNA video content should incorporate a combination of visual and verbal cues to create a more engaging and immersive experience for viewers. More realistic, authentic, and particularly character-specific voiceover could be utilized in VI SNA video content to attract more viewers. Furthermore, investing in the state-of-the-art technologies that can enhance human touch in consumer-brand interactions may be highly beneficial. As merely implementing voice as a human cue might not be sufficient according to the findings of this study, businesses are urged to experiment with various methods to convey human-likeness more effectively.

1. Acknowledgement

Not applicable.

2. Ethics and consent

Not applicable.

3. Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

4. Conflicting interests

Not applicable.

5. Funding

This research was supported by a Korea University Grant.

6. Authors' contributions

JW and NH developed research ideas and designed methodology. JW collected and analyzed the data, and drafted the initial manuscript. NH supervised the entire research process and contributed to the manuscript writing. YJ gave advice to the manuscript writing and reviewed the manuscript. All authors read and approved the final manuscript.

7. Authors' information

Jeeweon Wee Dept. of Human Ecology & Interdisciplinary Program in Sustainable Living System, Korea University Graduate School, Doctoral Student

Namhee Yoon Human Ecology Research Center, Korea University, Research Professor

Yoon-Jung Lee Dept. of Home Economics Education, Korea University, Professor

References

- Ahn, R. J., Cho, S. Y., & Sunny Tsai, W. (2022). Demystifying computer-generated imagery (CGI) influencers: The effect of perceived anthropomorphism and social presence on brand outcomes. *Journal of Interactive Advertising*, 22(3), 327–335. <https://doi.org/10.1080/15252019.2022.2111242>
- Arsenyan, J., & Mirowska, A. (2021). Almost human? A comparative case study on the social media presence of virtual influencers. *International Journal of Human-Computer Studies*, 155, 102694. <https://doi.org/10.1016/j.ijhcs.2021.102694>
- Bailenson, J. N. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. WW Norton & Company.
- Bailenson, J. N., Yee, N., Merget, D., & Schroeder, R. (2006). The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence: Teleoperators and Virtual Environments*, 15(4), 359–372. <https://doi.org/10.1162/pres.15.4.359>
- Banks, J., & Bowman, N. D. (2016). Avatars are (sometimes) people too: Linguistic indicators of parasocial and social ties in player-avatar relationships. *New Media & Society*, 18(7), 1257–1276. <http://dx.doi.org/10.1177/1461444814554898>
- Bartneck, C., Kanda, T., Ishiguro, H., & Hagita, N. (2007, August 26-29). *Is the uncanny valley an uncanny cliff?* [Conference session]. The 16th IEEE International Symposium on Robot and Human Interactive Communication, Jeju, Republic of Korea. <http://dx.doi.org/10.1109/ROMAN.2007.4415111>
- Bartneck, C., Kanda, T., Ishiguro, H., & Hagita, N. (2009, September 27-October 2). *My robotic doppelgänger: A critical look at the uncanny valley* [Conference session]. The 18th IEEE International Symposium on Robot and Human Interactive Communication, Toyama, Japan. <http://dx.doi.org/10.1109/ROMAN.2009.5326351>
- Block, E., & Lovegrove, R. (2021). Discordant storytelling, 'honest fakery', identity peddling: How uncanny CGI characters are jamming public relations and influencer practices. *Public Relations Inquiry*, 10(3), 265–293. <https://doi.org/10.1177/2046147X211026936>
- Brenton, H., Gillies, M., Ballin, D., & Chatting, D. (2005, September 5-9). *The uncanny valley: Does it exist* [Conference session]. Workshop on human-animated characters interaction at the 19th British HCI group annual conference: The bigger picture, Edinburgh, United Kingdom. <http://www.davidchatting.com/>
- Cheetham, M., Suter, P., & Jäncke, L. (2011). The human likeness dimension of the “uncanny valley hypothesis”: Behavioral and functional MRI findings. *Frontiers in Human Neuroscience*, 5, 126. <https://doi.org/10.3389/fnhum.2011.00126>
- Cheng, X., Bao, Y., Zarifis, A., Gong, W., & Mou, J. (2022). Exploring consumers' response to text-based chatbots in e-commerce: The moderating role of task complexity and chatbot disclosure. *Internet Research*, 32(2), 496–517. <https://doi.org/10.1108/INTR-08-2020-0460>
- Chidiac, D., & Bowden, J. (2023). When media matters: The role of media richness and naturalness on purchase intentions within influencer marketing. *Journal of Strategic Marketing*, 31(6), 1178–1198. <https://doi.org/10.1080/0965254X.2022.2062037>
- Cho, E., Molina, M. D., & Wang, J. (2019). The effects of modality, device, and task differences on perceived human likeness of voice-activated virtual assistants. *Cyberpsychology, Behavior and Social Networking*, 22(8), 515–520. <https://doi.org/10.1089/cyber.2019.0038>
- Choudhry, A., Han, J., Xu, X., & Huang, Y. (2022). I felt a little crazy following a 'doll': Investigating real influence of virtual influencers on their followers. *Proceedings of the ACM on Human-Computer Interaction*, 6(GROUP), 1–28. <https://doi.org/10.1145/3492862>
- Dabiran, E., Farivar, S., Wang, F., & Grant, G. G. (2023, June 11-16). *Engaging with virtual influencers: A netnography study* [Conference session]. European Conference on Information Systems, Kristiansand, Norway. <https://aisel.net>

- aisnet.org/ecis2023_rip/41
- DataBridge. (2022). *Global influencer marketing platform market: Industry trends and forecast to 2031*. DataBridge Market Research. <https://www.databridgemarketresearch.com/reports/global-influencer-marketing-platform-market>
- de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, 22(3), 331–349. <https://doi.org/10.1037/xap0000087>
- Ding, Y., Prepin, K., Huang, J., Pelachaud, C., & Artieres, T. (2014, May 5-9). *Laughter animation synthesis* [Conference session]. The 13th International Conference on Autonomous Agents and Multiagent Systems, Paris, France. <https://dl.acm.org/doi/10.5555/2615731.2615856>
- Dwivedi, Y. K., Ismagilova, E., Rana, N. P., & Raman, R. (2023). Social media adoption, usage and impact in business-to-business (B2B) context: A state-of-the-art literature review. *Information Systems Frontiers*, 25, 971–993. <https://doi.org/10.1007/s10796-021-10106-y>
- Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, 132, 138–161. <https://doi.org/10.1016/j.ijhcs.2019.07.009>
- Fernandes, T., & Oliveira, E. (2021). Understanding consumers' acceptance of automated technologies in service encounters: Drivers of digital voice assistants adoption. *Journal of Business Research*, 122, 180–191. <https://doi.org/10.1016/j.jbusres.2020.08.058>
- Finkel, M., & Krämer, N. C. (2022). Humanoid robots-Artificial. Human-like. Credible? Empirical comparisons of source credibility attributions between humans, humanoid robots, and non-human-like devices. *International Journal of Social Robotics*, 14(6), 1397–1411. <https://doi.org/10.1007/s12369-022-00879-w>
- Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007, August 22-25). *How interface agents affect interaction between humans and computers* [Conference session]. The 2007 Conference on Designing Pleasurable Products and Interfaces, Helsinki, Finland. <https://doi.org/10.1145/1314161.1314180>
- Franke, C., Groeppel-Klein, A., & Müller, K. (2023). Consumers' responses to virtual influencers as advertising endorsers: Novel and effective or uncanny and deceiving? *Journal of Advertising*, 52(4), 523–539. <https://doi.org/10.1080/00913367.2022.2154721>
- Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21(1), 1–31. <https://doi.org/10.1086/209380>
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007, September 17-19). *Creating rapport with virtual agents* [Conference session]. The 7th International Conference on Intelligent Virtual Agents, Paris, France. https://doi.org/10.1007/978-3-540-74997-4_12
- Griskevicius, V., Shiota, M. N., & Neufeld, S. L. (2010). Influence of different positive emotions on persuasion processing: A functional evolutionary approach. *Emotion*, 10(2), 190–206. <https://doi.org/10.1037/a0018421>
- Guadagno, R. E., Blascovich, J., Bailenson, J. N., & McCall, C. (2007). Virtual humans and persuasion: The effects of agency and realistic behavior. *Media Psychology*, 10(1), 1–22. <https://doi.org/10.1080/15213260701300865>
- Ham, J., Li, S., Shah, P., & Eastin, M. S. (2023). Are virtual influencers more effective in real-world advertisement settings? The impact of product type and social cues on consumers' perceptions and behaviors. *Journal of Advertising Research*, 64(3), 353–371. <http://doi.org/10.2501/JAR-2024-015>
- Ho, C.-C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, 26(6), 1508–1518. <https://doi.org/10.1016/j.chb.2010.05.015>
- Ho, C.-C., MacDorman, K. F., & Pramono, Z. A. (2008, March 12-15). *Human emotion and the uncanny valley: A GLM, MDS, and Isomap analysis of robot video ratings* [Conference session]. The 3rd ACM/IEEE International Conference on Human-Robot Interaction, Amsterdam, Netherlands. <http://dx.doi.org/10.1145/1349822.1349845>
- Ischen, C., Araujo, T. B., Voorveld, H. A., Van Noort, G., & Smit, E. G. (2022). Is voice really persuasive? The influence of modality in virtual assistant interactions and two alternative explanations. *Internet Research*, 32(7), 402–425. <https://doi.org/10.1108/INTR-03-2022-0160>
- Kim, I., Ki, C. W., Lee, H., & Kim, Y. K. (2024). Virtual influencer marketing: Evaluating the influence of virtual influencers' form realism and behavioral realism on consumer ambivalence and marketing performance. *Journal of Business Research*, 176, 114611. <https://doi.org/10.1016/j.jbusres.2024.114611>
- Koles, B., Audrezet, A., Moulard, J. G., Ameen, N., & McKenna, B. (2024). The authentic virtual influencer: Authenticity manifestations in the metaverse. *Journal of Business Research*, 170, 114325. [- 1208 -](https://doi.org/10.1016/</p></div><div data-bbox=)

- j.jbusres.2023.114325
- Lee, C. T., Pan, L. Y., & Hsieh, S. H. (2021). Artificial intelligent chatbots as brand promoters: A two-stage structural equation modeling-artificial neural network approach. *Internet Research*, 32(4), 1329–1356. <https://doi.org/10.1108/INTR-01-2021-0030>
- Li, X., Wu, C., Mai, F., & Jung, I. M. (2019). The effect of online reviews on product sales: A joint sentiment-topic analysis. *Information & Management*, 56(2), 172–184. <https://doi.org/10.1016/j.im.2018.04.007>
- Lim, R. E., & Lee, S. Y. (2023). “You are a virtual influencer!”: Understanding the impact of origin disclosure and emotional narratives on parasocial relationships and virtual influencer credibility. *Computers in Human Behavior*, 148, 107897. <https://doi.org/10.1016/j.chb.2023.107897>
- Longoni, C., Bonezzi, A., & Morewedge, C. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4), 629–650. <https://doi.org/10.1093/jcr/ucz013>
- Loosen, W., Häring, M., Kurtanović, Z., Merten, L., Reimer, J., van Roesel, L., & Maalej, W. (2018). Making sense of user comments: Identifying journalists’ requirements for a comment analysis framework. *SCM Studies in Communication and Media*, 6(4), 333–364. <http://dx.doi.org/10.5771/2192-4007-2017-4-333>
- Lou, C., Kiew, S. T. J., Chen, T., Lee, T. Y. M., & Ong, J. E. C. (2023). Authentically fake? How consumers respond to the influence of virtual influencers. *Journal of Advertising*, 52(4), 540–557. <https://doi.org/10.1080/00913367.2022.2149641>
- Louwerse, M. M., Graesser, A. C., Lu, S., & Mitchell, H. H. (2005). Social cues in animated conversational agents. *Applied Cognitive Psychology*, 19(6), 693–704. <https://doi.org/10.1002/acp.1117>
- Lucini, F. R., Tonetto, L. M., Fogliatto, F. S., & Anzanello, M. J. (2020). Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. *Journal of Air Transport Management*, 83, 101760. <https://doi.org/10.1016/j.jairtraman.2019.101760>
- MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies, Social Behaviour and Communication in Biological and Artificial Systems*, 7(3), 297–337. <https://doi.org/10.1075/is.7.3.03mac>
- MacDorman, K. F., Green, R. D., Ho, C. C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior*, 25(3), 695–710. <https://doi.org/10.1016/j.chb.2008.12.026>
- Martín, L. R., Martín, I. R., & Sastre, D. M. (2021). Los influencers virtuales como herramienta publicitaria en la promoción de marcas y productos. Estudio de la actividad comercial de Lil Miquela. *Revista Latina de Comunicación Social*, 79, 69–90. <https://doi.org/10.4185/RLCS-2021-1521>
- Mathur, M., & Reichling, D. (2016). Navigating a social world with robot partners: A quantitative cartography of the uncanny valley. *Cognition*, 146, 22–32. <https://doi.org/10.1016/j.cognition.2015.09.008>
- Miao, F., Kozlenkova, I. V., Wang, H., Xie, T., & Palmatier, R. W. (2022). An emerging theory of avatar marketing. *Journal of Marketing*, 86(1), 67–90. <https://doi.org/10.1177/0022242921996646>
- Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2), 98–100. <http://dx.doi.org/10.1109/MRA.2012.2192811>
- Moriuchi, E. (2019). Okay, Google!: An empirical study on voice assistants on consumer engagement and loyalty. *Psychology and Marketing*, 36(5), 489–501. <https://doi.org/10.1002/mar.21192>
- Nowak, K. L., & Biocca, F. (2003). The effect of agency and anthropomorphism on users’ sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 12(5), 481–494. <https://doi.org/10.1162/105474603322761289>
- Nowak, K. L., & Fox, J. (2018). Avatars and computer-mediated communication: A review of the definitions, uses, and effects of digital representations. *Review of Communication Research*, 6, 30–53. <https://doi.org/10.12840/issn.2255-4165.2018.06.01.015>
- Parmar, D., Olafsson, S., Utami, D., Murali, P., & Bickmore, T. (2022). Designing empathic virtual agents: Manipulating animation, voice, rendering, and empathy to create persuasive agents. *Autonomous Agents and Multi-agent Systems*, 36(1), 17. <https://doi.org/10.1007/s10458-021-09539-1>
- Pavlidou, K. (2021). *The uncanny valley: The humanlikeness of chatbots and its paradoxical impact on consumers’ purchase intention in e-commerce* [Unpublished master’s thesis]. Tilburg University.
- Pelau, C., Dabija, D.-C., & Ene, I. (2021). What makes an AI device human-like? The role of interaction quality, empathy, and perceived psychological anthropomorphic characters in the acceptance of artificial intelligence in the service industry. *Computers in Human Behavior*, 122, 106855. <https://doi.org/10.1016/j.chb.2021.106855>
- Petty, R. E., Fabrigar, L. R., & Wegener, D. T. (2003). Emotional factors in attitudes and persuasion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 752–772). Oxford University Press.

- Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13(4), 707–727. <https://doi.org/10.1017/S0140525X00081061>
- Pittman, M., & Reich, B. (2016). Social media and loneliness: Why an Instagram picture may be worth more than a thousand Twitter words. *Computers in Human Behavior*, 62, 155–167. <https://doi.org/10.1016/j.chb.2016.03.084>
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.
- Rhee, C. E., & Choi, J. (2020). Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Computers in Human Behavior*, 109, 106359. <https://doi.org/10.1016/j.chb.2020.106359>
- Robinson, B. (2020). Towards an ontology and ethics of virtual influencers. *Australasian Journal of Information Systems*, 24, 333–345. <https://doi.org/10.3127/ajis.v24i0.2807>
- Schroeder, J., & Epley, N. (2016). Mistaking minds and machines: How speech affects dehumanization and anthropomorphism. *Journal of Experimental Psychology: General*, 145(11), 1427–1437. <https://doi.org/10.1037/xge0000214>
- Sivaramakrishnan, S., Wan, F., & Tang, Z. (2007). Giving an “e-human touch” to e-tailing: The moderating roles of static information quantity and consumption motive in the effectiveness of an anthropomorphic information agent. *Journal of Interactive Marketing*, 21(1), 60–75. <https://doi.org/10.1002/dir.20075>
- Slater, M., & Banakou, D. (2021). The golden rule as a paradigm for fostering prosocial behavior with virtual reality. *Current Directions in Psychological Science*, 30(6), 503–509. <https://doi.org/10.1177/09637214211046954>
- Van Zant, A. B., & Berger, J. (2020). How the voice persuades. *Journal of Personality and Social Psychology*, 118(4), 661–682. <https://psycnet.apa.org/doi/10.1037/pspi0000193>
- Wang, S., Lilienfeld, S. O., & Roachat, P. (2015). The uncanny valley: Existence and explanations. *Review of General Psychology*, 19(4), 393–407. <https://doi.org/10.1037/gpr0000056>
- Warren, C., Batra, R., Loureiro, S. M. C., & Bagozzi, R. P. (2019). Brand coolness. *Journal of Marketing*, 83(5), 36–56. <https://doi.org/10.1177/0022242919857698>
- Wilson, T., Meyers, J., & Gilbert, D. (2001). Lessons from the past: Do people learn from experience that emotional reactions are short-lived? *Personality and Social Psychology Bulletin*, 27(12), 1648–1661. <https://doi.org/10.1177/01461672012712008>
- Yu, J., Dickinger, A., So, K. K. F., & Egger, R. (2024). Artificial intelligence-generated virtual influencer: Examining the effects of emotional display on user engagement. *Journal of Retailing and Consumer Services*, 76, 103560. <https://doi.org/10.1016/j.jretconser.2023.103560>