"I Like the Sound of That": Understanding the Effectiveness of Audio in Sports Ads

Journal:	*Internet Research*
Manuscript ID	INTR-10-2023-0898.R2
Manuscript Type:	Research Paper
Keywords:	Super bowl, Advertising, Music analytics, Machine learning, Voice analytics

Manuscript ID INTR-10-2023-0898.R1
"I Like the Sound of That": Understanding the Effectiveness of Audio in Sports Ads

We would like to thank the editor and reviewers for their time in evaluating our manuscript. Reviewers 1 and 2 accepted the paper with no further comments. However, reviewer 3 recommended some minor revisions and the editorial office requested some changes. Therefore, we have revised the paper accordingly as follows.

Comment	**Response**
Reviewer #3 This manuscript has been substantially revised and improved in response to the review comments. The revisions show that the authors have taken the review comments seriously and have made improvements accordingly.	Thank you for your helpful insights that have helped us to improve the paper and your positive response to the revised manuscript.
I would like to summarize the changes made by the author as follows. 1. In terms of organization, the authors deleted the research results from the introduction as suggested, and explained the concept of signal sender and receiver more clearly, which improved the logic and readability of the paper.	Many thanks for your positive assessment of the revised organization and readability of the paper.
2. The authors increased the transparency of the study by adding a description of the sample exclusion criteria. Meanwhile, in the introduction section, the motivation and significance of the study are more clearly stated to emphasize the value of the study.	Thank you for your positive comments on the sample exclusion criteria and motivational significance that have been added to the revised paper.
3. Some concepts such as "voice frames" are defined as necessary, and literature support is added to some of the thesis statements to enhance the rigor of the argumentation.	Thank you for recognizing these recommended changes.
4. A discussion of cognitive consistency theory is included in the literature review section to enrich the theoretical foundation. In the Future Research section, the possible roles of other appeal strategies are explored to broaden the perspective of the study.	Thanks for these suggestions and the subsequent improvements that have been made.
5. Streamlined and reorganized the Discussion and Conclusion sections, deleted repetitive statements, and made the paper more concise.	Thank your positive assessment of these changes.
6. The first paragraph of the Discussion section was rewritten to highlight the main findings of the study and to make the contribution of the study more distinct.	We value your evaluation of our revisions.
Overall, the quality of the manuscript has been greatly improved. However, there are some suggestions for the authors to consider as follows. 1. The application of signaling theory is not sufficient. It is suggested that further elaboration on how sound and music features as signals affect consumer	Thank you for this suggestion. A further elaboration has been provided, extending the content of the signaling theory subsection.

perception and evaluation is needed to highlight the value of this theoretical perspective.	
2. Although some suggestions on audio advertisement design are mentioned in the section of practical revelations, they can be more specific, such as giving some quantitative reference values, so as to enhance the operability of the revelations.	Thank you for this suggestion. Quantitative reference values have now been added to the practical implications section.
3. The language can be more concise, for example, some long sentences can be considered broken to improve the clarity of expression.	Thank you for this suggestion. Some longer sentences have now been broken into shorter sentences.
All in all, the authors showed a rigorous and responsible scientific research attitude in the revision process, accepted the reviewers' comments with an open mind, and made serious revisions.	Thank you for your positive assessment of our revisions.
Editorial Office The manuscript has room for improvement before it is accepted for publication. Below are some examples. 1. There are some typos (e.g., "March 3rd, 2022" in the reference list should be "March 3, 2022", "MacKenzie" vs. "Mackenzie", "christmas" in the reference list should be "Christmas", a space is missing in "16calculated"). The spelling of some words should be changed because the manuscript uses American English (e.g., "towards" should be "toward", "emphasise", "Humour Appeal" in Figure 1 and Table 3).	Thank you for these suggestions. We have made these amendments, as requested.
2. The manuscript should use the required reference format and citation style (see the Author Guidelines at https://www.emeraldgrouppublishing.com/journal/intr). Remove the spaces between the initials of each author or editor name in references (e.g., "Anikin, A. S." should be "Anikin, A.S.", etc.). Author names should be properly initialized in references (e.g., "Belch, George E., and Michael A. Belch.", "Villarroel, Ordenes, F."). All ", &" between author names in references should be "and" without the comma (e.g., "Fürst, A., Pečornik, N., & Binder, C." should be "Fürst, A., Pečornik, N. and Binder, C.", etc.). All "&" between author names in citations should be "and" instead (e.g., "Van Zant & Berger, 2020"). A paper title should appear within a pair of double quotation marks. Volume, issue (if any) and page numbers should follow the words "Vol.", "No." and "pp." respectively. Different parts of a reference should be joined together with commas (e.g., after the publication year, after the paper title). References should be correct (e.g., some with author names that start with "G" are mixed	Thank you for these suggestions. We have revised the references and citations accordingly.

together). A reference from an electronic source must include both the URL and the access date after the words "available at:" and "accessed" respectively (e.g., Harting, 2020; Statista, 2021). References should be sorted by author name(s) and year (e.g., "Hofacker, C. F..." should appear before "Holbrook, M. B...", "Kao, K. C..." should appear after "Kanner, B...", "Knoeferle, K. M..." should appear before "Ko, S. J...", "Lowe, M. L..." should appear before "Ludwig, S...", "Raithel, S..." should appear before "Raudenbush, S.W...", "Tannenbaum, M. B..." should appear before "Tavassoli, N. T..."). Citations should use the appropriate style according to the context (e.g., "Mackenzied et al., (1986, p. 130)" should be "Mackenzie et al. (1986, p. 130)" without the comma after "et al."). Citations should be correct (e.g., "Martín-Santana et al., 2015" should be either "Martín-Santana et al., 2015a" or "Martín-Santana et al., 2015b"). References not cited should be removed (e.g., Herhausen et al., 2020; Styvén et al., 2020). Some references are not cited (e.g., Rázuri, 2015) while some citations cannot be found in the reference list (e.g., "Rázuri et al. (2015)").	
3. The manuscript should be consistent. Some p-values include a zero before the decimal point while some do not. The construct "The median intensity of voiced frames in an ad" in H1 becomes "Mean Intensity" in Figure 1. The construct "the liking behavior of an ad" becomes "Propensity to Like Ad" in Figure 1. The construct "The proportion of voice frames in an ad" in H4 becomes "% Voice Frames" in Figure 1 and "Percent of Voice Frames" in Table 1. The control variable should be either "Celebrity ad appeals" or "Celebrity appeals" or "Celebrity appeal" in the main text and in Figure 1. The step "Variable Preparation" in Figure 2 should start with a verb like other steps.	Thank you for these helpful suggestions. We have made the requested changes to the paper.
4. The manuscript needs to be carefully checked. The main text says that "The first step in the research process was to select a sample of Super Bowl ads for analysis... during the period 2000 to 2020" while the first step in Figure 2 is "Select Sample of Superbowl Ads (top-10, 2000-21)" (in which "top-10" should not be hyphenated). The main text says that "a mean tempo-duration interaction of 4999.66" while the value in Table 1 is "4699.66". The main text says that "In terms of our control variables... a celebrity; 32% of the sample included a fear appeal..." (in which the semicolon should be a comma instead) while the percentage in Table 1 is "30.4%".	Thanks for these recommendations. The paper has been amended accordingly.

5. There are ways to improve the readability of the manuscript (e.g., "In obverse to" should be "In contrast to" instead, both "congruency" may be "congruence" instead, literature review and conceptual model should be broken in to two sections and the last paragraph of the introduction section should be revised accordingly, the hypotheses should use the present tense rather than the future tense, individual control variables should be "appeal", the axis labels and the keys in the charts in Figure 5 should use the upper case in the first letter of the first word, the main text to report explicitly whether H3b and H5b are supported or not respectively, "First" in the "Limitations and future research" section should be followed by "Second" and so on, etc.). Keywords in the system and in the manuscript should be identical. Figure 1 should be revised to emphasize the moderating effect in H5b ("The tempo of music during an ad will positively moderate the relationship between duration and the liking behavior of an ad"). Besides, it is unusual to propose a moderating effect of a construct on a relationship without first proposing the direct relationship.	Thank you for these suggestions. They have been taken onboard in the revised manuscript. Furthermore, we replaced the “moderation effect” with “interaction effect”.
The manuscript should present items in the same order across tables, figures and in the main text (e.g., the footnote explaining "dom", "autocor" and "final" in Figures 3 and 4 should be sorted in the same order as they are in the keys, variables in Tables 1 and 2 should appear in the same order as they are in the hypotheses). Words in some figures are too small to be seen clearly. Words in the chart on the left in Figure 6 should not overlap with plots in the chart. Figures and tables should be self-explanatory.	Thank you for these suggestions. They have assisted us in revising the paper further.
The y-axis in the charts in Figures 3 and 4 should include the unit (i.e., "kHz"). Check the unit of the x-axis label "ms" in Figures 3 and 4 are correct. The part "-90" in Figures 3 and 4 should be explained by including a unit after the number or an explanation in the footnote. Some variables in Table 1 should be followed by the unit.	The figures and tables have been revised accordingly.
Sentences describing prior studies should use the past tense (e.g., "Chung and Kalnins (2001) explain", "Martín-Santana et al. (2015a) and Zander (2006) propose", "Biswas et al. (2019) claim", "Lowe and Haws (2017) reveal", "Kim and Zauberman (2019) discover", "Taylor (2016) discusses", , "Martín-Santana et al. (2015a) and Zander (2006) propose", "Van Zant and Berger (2020) also discover", "Kamiloğlu et al.	These have now been changed to the past tense.

(2020) infer", "Mackenzie et al., (1986, p. 130) define").	
6. Acknowledge the conference paper upon which this manuscript is developed.	The conference presentation has now been added as an acknowledgement.
Please revise the manuscript thoroughly and submit the revised version. Thank you.	Thank you for your suggestions. We have revised the paper.

"I Like the Sound of That":

Understanding the Effectiveness of Audio in Ads

Abstract

Purpose: Sports advertisements such as the Super Bowl showcase products and brands that have invested increasingly large sums financially to gain viewers' attention. However, how audio features in advertisements impact viewers' behavior remains unexplored.

Design/methodology/approach: Using the lens of signaling theory, this research uses advanced data analytics of voice and music audio in Super Bowl advertisements to examine its impacts on viewers.

Findings: Results show that advertisement viewers prefer more voiced frames and have a greater liking behavior of voiced frames with a low intensity (less loud) and a higher F1 frequency, which is typically associated with male vocal tracts. A fast music tempo works better for longer advertisements. The model controls for various types of ad appeals. The research underlines the importance of voice and music in signaling successful brand features that are likely to increase the ad-liking behavior of consumers (positive effect).

Research implications: The current research implies that brands advertising through sports ads must carefully select voice actors and music in order to provide the most positive signals for a brand to have the most significant effect and, thus, a greater return on the high sums invested in the ads.

Keywords: machine learning, voice analytics, music analytics, Super Bowl, advertising.

INTRODUCTION

Brands have invested increasingly large sums into sports advertisements in recent years, including for the Olympics and World Cup soccer, and spearheaded in the U.S. by the Super Bowl. The Super Bowl is a leading sporting event in the U.S. and the most-watched televised event on U.S. television. Large corporations leverage Super Bowl advertising to enhance the awareness and visibility of their products and services, driven by high viewership and broad demographics of viewers (Kanner, 2004). Advertising is tremendously intertwined with the Super Bowl – before, during, and after the event – and commercials have become an integral part of it. A considerable amount of research has examined Super Bowl ads from a variety of perspectives, such as advertising effectiveness through elements such as ad likability, ad recall, and buzz (Shan *et al.,* 2023; Raithel *et al.,* 2016; Newell and Henderson, 1998), the economic value of ads in terms of stock market returns, financial returns and market value (Lee and Ko, 2021; Kim *et al.,* 2015; Tomkovick *et al.,* 2011; Eastman *et al.,* 2010; Fehle *et al.,* 2005; Choong *et al.,* 2003; Kim and Morris, 2003) the impact of ads on the online product or brand searches (Chandrasekaran *et al.,* 2018), the role of word-of-mouth conversations and media coverage concerning ad persuasiveness (Nail, 2007), and engagement with brands via social media (Noh *et al.,* 2021). However, the research has not yet explored the unique audio elements in the body of advertisements – such as voice and music features –that are a potential source of influence for viewers. Such features or signals embedded into ads are processed automatically by viewers.

Audio data analytics development allows a platform for investigating the effect of audio features on consumers (Hildebrand *et al.,* 2020). Indeed, current research on human coding has admitted its limitations and encouraged more research based on machine coding of video data (Tellis *et al.,* 2019). Empirical research utilizing features from real audio in advertisements is scarce. Audio analytics, offering huge research potential (Dawar, 2018), is used by very few studies in advertising. Most existing advertising studies focus on *the use of experiments and field studies* (Huang and Labroo, 2020; Rodero *et al.,* 2017; Martín-Santana *et al.,* 2015a; Chattopadhyay *et al.,* 2003) rather than the integration of unstructured features from the *actual behavior data* using analytics. Furthermore, there is limited advertising research investigating the effect of voice, music, and ad appeals on consumer outcomes using data analytics techniques. Understanding these elements could unlock more value for companies through more powerful advertisements that produce a greater return on investment in the increasingly expensive and crowded sports ads space, where companies vie for attention. Considering the importance of audio features embedded in ads, there is a pertinent theoretical and managerial need for insight into how voice and music features influence consumers' ad-liking behavior (e.g., a Super Bowl ad).

The current work identifies the critical voice and music features relevant to ads to address the research gap. We examine the effects of various voice and music features on consumers' liking behavior of an ad. Specifically, we use YouTube video data from the most frequently advertised products and brands in Super Bowls from 2000 to 2020. Data from Google (2016) suggests that YouTube is the critical channel for accessing Super Bowl advertisements that have been missed from the live broadcast; 48% of viewers accessed missed advertisements on YouTube during the game. Consumers spent 641,000 hours watching Super Bowl ads on YouTube in 2019 (Statista, 2021). Given the popularity of YouTube videos and viewer engagement, Super Bowl ads represent an insightful research context.

Noting that sound features that the firm controls can function as signals (Biswas *et al.,* 2019), we propose that firms (i.e., signal senders) can signal the sound and music features to consumers (i.e., signal receivers) to influence their behavior (Spence, 1973) and evaluations (Guha *et al.,* 2022). Signaling theory has been applied in numerous studies, such as voice assistants (Guha *et al.,* 2023; Jain *et al.,* 2022) and e-commerce (Guo *et al.,* 2020). We draw from signaling theory to propose the following contributions. First, the study develops and applies a research process integrating advanced data analytics methods to test a research model for a number of existing and new theoretical relationships between voice, music, and ad-liking behavior. Research methods for integrating audio analytics in behavioral research have yet to be developed, and we provide a clear blueprint for such research. Hofacker *et al.* (2016) posited the power and rich insights that data analytics can offer to our understanding of consumer behavior. This view is echoed by Zhang and Watson (2020), whereby firms can precisely understand customers' evolving preferences and design suitable engagement tactics. Data analytics becomes the bridge to a more informed customer behavior in today's data-driven environment. Second, conceptualizing voice and music features as signals, we demonstrate that these features affect consumer perceptions of advertisements by analyzing actual audio features extracted from ads combined with *actual data on consumer interaction with the videos* and highlighting which sound and music features exert the most important effects. Third, this work also reveals the significant interaction effect between sound and music features. While prior research did not generate conclusive results (Martín-Santana *et al.,* 2015a; Zander, 2006), our research shows that a high level of fit between sound and music features leads to an increasing level of liking of an ad. Finally, we demonstrate that firms can design effective audio advertisements by conceptualizing voice and music features as signals.

The structure of this paper is as follows. The next section examines the theoretical background of signaling theory and audio features in advertisements, followed by a section on developing research hypotheses. The fourth section provides a step-by-step explanation of the original research process applied to extract and analyze the audio and other features examined in the research. The fifth section details the statistical analysis results to test the research model and hypotheses. Finally, the last section discusses the results, with implications for further research and practice, limitations of the study, and conclusions.

LITERATURE REVIEW

Attitude toward advertisement

MacKenzie *et al.* (1986, p.130) defined attitude toward the ad as the "predisposition to respond favorably or unfavorably to a particular advertising stimulus during a particular exposure situation." The content of the ad, including brand attributes, headlines, images, music, and celebrities, all contribute to the ad's formation (Oakes, 2007; Biehal *et al.,* 1992). Furthermore, attitude toward the ad does not necessarily lead to a favorable brand attitude, especially when one brand is superior to another. However, the relationship between ad attitude and brand choice seems direct for two or more similar brands. Therefore, our focus is on consumers' liking behavior and does not consider whether one brand is superior to others.

Signaling theory

Signaling theory, first formulated by Spence (1973) in the field of labor economics, has been used extensively in management research to describe the behaviors of two parties when each has access to different information, particularly in strategy, human resource management, and entrepreneurship (Connelly *et al.,* 2011). Furthermore, signaling theory has also been applied to a certain extent for consumer advertising. Chung and Kalnins (2001) explained that advertising, price, and location are the key signals to consumers from firms. They also find that signaling reduces consumers' search costs but warns against false signals and the marginal utility of too many signals. Advertisements can signal product quality (Thomas *et al.,* 1999) and even encourage a willingness to pay for both high- and low-quality products (Tsui, 2012). In addition, signals in ads are persuasive enough to increase willingness to pay and intention to buy green products (Sun *et al.,* 2020; Berger, 2019). More recently, voice assistants' speech has been a signal that affects consumers' evaluation of their artificiality and intelligence.

Firms (i.e., signal senders) may develop advertisements that contain various types of signals, which can be verbal and nonverbal. Exemplar signals are visual imagery, dialogue, narrative, and the inclusion of various ad appeals, such as celebrities, humor, sex, or danger. Another type of signal is audio, with speakers' voices and the inclusion of music in an ad. Nonverbal communication in voices influences individuals' judgment (Van Zant and Berger, 2020). Consumers rely on nonverbal cues such as loudness, pitch, and frequency to appraise a speaker's emotion (Juslin and Laukka, 2003), an essential carrier of human emotions. Voices contain many different features that provide non-verbal signals regarding the brand being advertised. Such voice characteristics may be assessed and interpreted automatically by listeners (i.e., signal receivers) (Hildebrand *et al.,* 2020).

Furthermore, music in advertisements serves not just as a background element but as a powerful signal that influences emotions and moods. The tempo, key, and genre of music can evoke specific feelings, from excitement and joy to nostalgia and calmness, setting the tone for the consumer's reception of the product or message (Tavassoli and Lee, 2003; Stewart and Punj, 1998). Music can align with the brand's identity, reinforcing the intended message and creating a memorable association in the consumer's mind (North *et al.,* 2016). Prior research on music also demonstrates that the background music, volume, tempo, and frequency influence the amount of time spent in retail stores, the amount purchased, evoke images and visual associations, and reduce the perceived distance between the sources and perceivers (Sunaga, 2018). Consumers' automatic processing of voices as non-verbal communication renders it particularly powerful for brands since they may encode audio features to provide a distinctive brand profile associated with audio signals. Brands that skillfully incorporate significant audio elements can create a deeper connection with their target audience, signaling an understanding and appreciation of their consumers' preferences.

CONCEPTUAL MODEL DEVELOPMENT

Voice intensity

Human speech generation is a function of both psychological and anatomical features. Regarding anatomical features, focusing on the human head, vocal utterances will depend upon aspects such as the size, shape, length of the tongue and teeth, and the size and shape of the nasal and oral cavities (Zhang, 2016). Together, these elements impact the extent to which the air stream is compressed and oscillates at various amplitudes and frequencies, creating an individual's unique voice pattern. Audio analytics concentrates on sound waves, which are typically described in four key dimensions: amplitude (intensity or loudness), frequency (or pitch), time (or duration), and spectrum (other complex wave components) (Sueur, 2018).

Voice is a type of audio and is capable of influencing consumer assessments, and "the choice of the voice in radio advertising is one of the most important decisions an advertiser is facing" (Martín-Santana *et al.,* 2015a). An important feature, the loudness of a signal, has been found to strongly influence individual emotions and consumer preferences. Biswas *et al.* (2019) claimed that low-volume ambient music increases relaxation levels, leading to a greater preference for healthy foods. In contrast, individuals perceive a sound increasing in loudness as an approaching object because such an object demands heightened vigilance due to the possibility of danger or threat (Bannister, 2020). Such findings align with Leongómez *et al.* (2021), indicating individuals may also modulate their voices strategically to signal nonverbal communication and characteristics such as threat and formidability. Moreover, louder voices tend to indicate aggression or anger (Juslin and Laukka, 2003). Louder voices signal greater dominance and extraversion (Tusing and Dillard, 2000; Scherer and Giles, 1979). Thus, the loudness of a voice is expected to influence consumers' preferences and liking of an advertisement.

Related to this, the variability in loudness can also influence individual behavior. Voice characteristics affect individuals' inferences of a speaker's social rank (Ko *et al.,* 2015). In particular, listeners tend to associate greater loudness variability with dominance, authority, and high social status. Van Zant and Berger (2020) also discovered that a more significant variation in loudness can manifest perceptions of confidence. Variability is associated with high-arousal emotions such as happiness (Juslin and Laukka, 2003). Kamiloğlu *et al.* (2020) inferred that happy voices exhibit more significant amplitude variability. Finally, voice frames refer to the individual frames of each video examined in the analysis in the Soundgen package. Thus, we propose the following hypotheses:

H1. The median intensity of voiced frames in an ad negatively relates to an ad's liking behavior.

H2. The standard deviation of intensity for voiced frames in an ad positively relates to the liking behavior of an ad.

The median intensity of voice frames will be used instead of the mean intensity of voice frames to avoid the potential of multicollinearity with the standard deviation of intensity for voiced frames.

The frequency of the voice

In addition to loudness intensity, frequency is another characteristic of voice that refers to the sound properties that determine pitch. Pitch and intensity are the two fundamental acoustic features of voice. Frequency refers to “the oscillations of sound waves, as measured in hertz (Hz)” (Hagtvedt and Brasel, 2016, p.552). F1 (around 500 Hz) and F2 (around 1500 Hz) represent the resonant frequencies of the vocal tract and are the important acoustic characteristics of a human voice (Kamiloğlu, *et al.,* 2020). Both F1 and F2 “determine the quality of vowels and are said to be sufficient to disambiguate the meaning in any given vowel” (Pathak *et al.,* 2017, p.577).

Pitch can be used in marketing communications to affect individual perceptions. Lowe and Haws (2017) revealed that pitch influences perceptions of product size. In contrast to a higher pitch, consumers associate a larger product size with a lower pitch in voice. Furthermore, high-pitched voices tend to be evaluated as more nervous and less trustworthy, with reduced force and potency (Apple *et al.,* 1979). On the obverse, politicians with lower-pitched voices were found to have an additional benefit in election campaigns (Tigue *et al.,* 2012) with more competence, trustworthiness, and persuasiveness (Guyer *et al.,* 2019; Oleszkiewicz *et al.,* 2017). Generally, dull, smooth, low-pitched voices that tend to be associated with males are typically more persuasive, boosting perceptions of the warmth of a speaker, behavioral intentions, and attitude toward them (Zoghaib, 2019). Voices that are lower in pitch also engender distinguishing, positive attitudes, associations, and greater brand recall, regardless of the speaker's gender (Zoghaib, 2017).

Finally, a human voice is one of the most important carriers of human emotions (Lortie *et al.,* 2017). Individuals rely on nonverbal cues such as pitch, loudness, and spectrum to evaluate a speaker’s emotion (Rázuri, 2015). Therefore, the greater the proportion of an ad that contains a human voice, the easier it is for listeners to assess the speaker’s emotion. Thus, the following hypotheses are proposed:

H3a. The mean F1 frequency for voiced frames in an ad positively relates to the liking behavior of an ad.

H3b. The mean F2 frequency for voiced frames in an ad negatively relates to the liking behavior of an ad.

H4. The proportion of voice frames in an ad positively relates to the liking behavior of an ad.

Congruence between music tempo and duration

A sensory signal can influence consumer inferences even when such cues are not directly diagnostic of the product. The background music tempo (i.e., beats per minute) is a typical example of such sensory signals (Zhu and Meyers-Levy, 2005). In physical stores, "the speed in which a [background] music progresses" influences the affective state of the listener (Spangenberg *et al.,* 2005). Music tempo may lead to a greater positive affective reaction. In contrast to slower-tempo music, music with a faster tempo stimulates heightened pleasure and arousal and thus influences greater impulse purchases (Oakes, 2003). In a restaurant context, fast music can manifest positive arousal, influencing consumers' expectations of taste for food and their intention to purchase (Pantoja and Borges, 2021). Additionally, fast-tempo music can increase physiological variables such as breathing, blood pressure, and heart rates (Vijayaraman and Ramadoss, 2019).

Notwithstanding the effects of music tempo, music tempo alone may not necessarily significantly impact an affective level (Feng, 2014; Caldwell and Hibbert, 2002). The effectiveness of music is linked to the congruence or musical fit with the advertising message (Martín-Santana *et al.,* 2015a; 2015b; Zander, 2006). The effects of music tempo on perceived duration (Caldwell and Hibbert, 2002; Kellaris and Kent, 1992) suggest potential positive outcomes of the congruence between music tempo and duration. To illustrate, Kim and Zauberman (2019) discovered that music with a higher number of beats per minute tends to increase the impatience of listeners as compared to music with a slower tempo. An explanation of this may be due to a potential relationship between a greater tempo and the perceived duration of waiting time (Oakes and North, 2008). Similarly, consumers perceive fast-tempo music's duration as longer than slow-tempo music (Caldwell and Hibbert, 2002). Such studies provide additional evidence that individuals associate fast tempo with long duration. From the cue congruence lens, a high degree of fit among stimulus characteristics often generates positive evaluations (Kao *et al.,* 2020; Purohit and Srivastava, 2001). Previous marketing research demonstrates a positive effect when multisensory cues are congruent, including congruence between sound and smell, sound and vision, sound and taste, and smell and touch (Knoeferle *et al.,* 2016; Krishna *et al.,* 2010). Congruence between multisensory cues has been well-evidenced. Music tempo and duration are examples of sensory cues, and a high degree of congruence between the two cues tends to generate positive evaluations (Fürst *et al.,* 2021; Oakes, 2007). Accordingly, it is posited that a fast tempo is effective only for lengthier advertisements because of the greater congruence. We hypothesize the following:

H5a. The tempo of music during an ad positively relates to the liking behavior of an ad.

H5b. There is a significant interaction effect between the tempo of music in an ad and the duration in predicting the liking behavior of an ad.

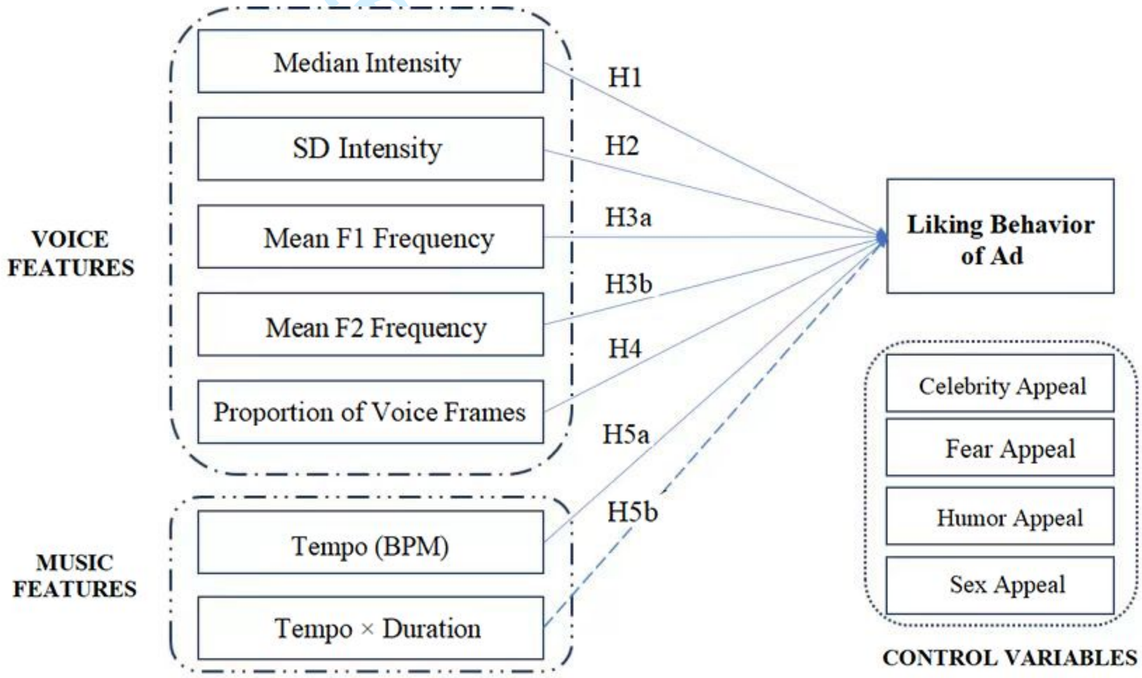
Control Variables: Advertising Appeals

Super Bowl ads have several variables that could affect the liking behavior of an ad that may need to be controlled for in our investigation. To illustrate, Taylor (2016) discussed the mixed effects of different ad appeals, such as humor and celebrities. An essential variable of an ad appeal is the appearance of a celebrity. Celebrity ad appeals are well known to influence positive outcomes, including message recall (Buttle *et al.,* 2000) and consumer persuasion (Belch and Belch, 2013). Building on the source-credibility theory, brand endorsers' expertise, trustworthiness, and attractiveness positively influence brand attitudes and brand credibility (Wang and Scheinbaum, 2018). Consumers are more willing to accept the arguments presented in advertisements when endorser credibility is high (Grewal *et al.,* 1994). A celebrity's physical attractiveness increases arousal, subsequently influencing consumers' information processing (Roozen and Claeys, 2010).

Other variables of ad appeals have also been researched extensively in the literature over recent decades. For example, emotional ad appeals are extremely important in advertising (Li, 2019; Poels and Dewitte, 2019; Holbrook and Batra, 1987). Fear appeals, including elements of danger or threat, have been extensively examined in the literature and have been found to have positive effects on attitudes, intentions, and behaviors (Tannenbaum *et al.,* 2015), but with varying effectiveness in different circumstances (Moraes *et al.,* 2019; Akram *et al.,* 2018; Antonetti *et al.,* 2015; Krisjanous *et al.,* 2013; Morales *et al.,* 2012). Sex or erotic appeals are another popular device applied to advertising and are widely used in areas such as personal care, perfume, apparel, and alcoholic beverages. They positively affect ad recognition and recall, but the same effect is not found for brand recognition and recall (Wirtz *et al.,* 2018). Sex appeals generally are found to have mixed effects on consumer outcomes such as liking and preferences (Wyllie *et al.,* 2014; Dahl *et al.,* 2009; Sengupta and Dahl, 2008; LaTour and Henthorne, 1994; Severn *et al.,* 1990). Humor is another popular appeal in advertising, including that of the

Super Bowl (Taylor, 2016). Humor in advertising is a positive driver of attitude, attention, consumption, and positive affect, but it has reduced credibility and mixed effectiveness in some studies (Hendriks and Strick, 2019; Eisend, 2009; De Pelsmacker and Geuens, 1999). Overall, this study controls four types of popular ad appeals: celebrity appeal, fear appeal, humor appeal, and sex appeal. The research model combines the above research hypotheses (see Figure 1); specifically, the current work examines three areas of features in an ad: voice features (H1 to H4) and music features (H5a and H5b), with ad appeals as control variables.

Figure 1. Research Model.

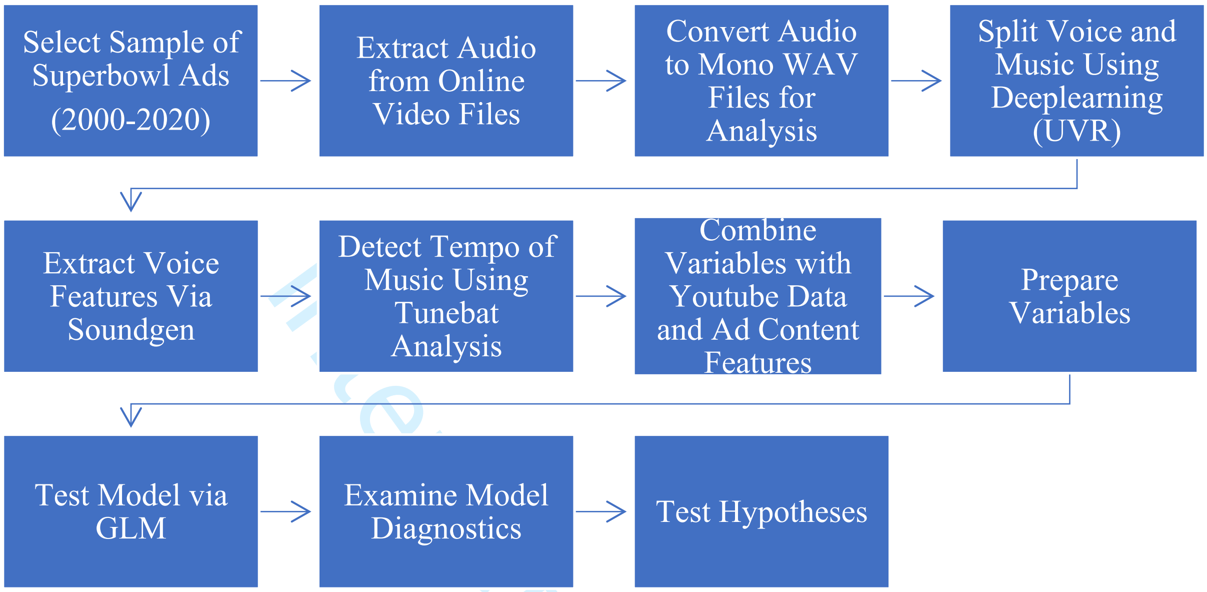


Note: A dashed line relationship indicates an interaction effect.

METHODOLOGY

The steps in the research process are outlined in Figure 2. The steps include selecting the sample of ads included in the study and extracting audio data, pre-processing the audio files for analysis, extracting audio features, merging data sets, data preparation, testing the research model, model diagnostics, and final testing and reporting of hypothesis results. Let us examine each of the steps in the process in turn.

Figure 2. Summary of Research Process.



Sample selection and data collection

The first step in the research process was to select a sample of Super Bowl ads for analysis. The investigation focused on ten brands with the most advertisements from 2000 to 2020: Budweiser, Bud Light, Doritos, Coca-Cola, eTrade, Kia, Pepsi, Hyundai, NFL, and Toyota[1]. This helped to focus the study on the most common audio signals in Super Bowl advertisements and manage the study's heavy computational requirements. As control variables, ads were coded according to whether they had specific ad appeals: presence of a celebrity, fear appeal, sex appeal, and humor appeal. The initial sample was 247 ads.

Data pre-processing: file conversion and audio splitting

[1] Data were originally sourced from Super Bowl Ads (http://www.superbowl-ads.com/) and compiled by FiveThirtyEight (https://github.com/fivethirtyeight/superbowl-ads).

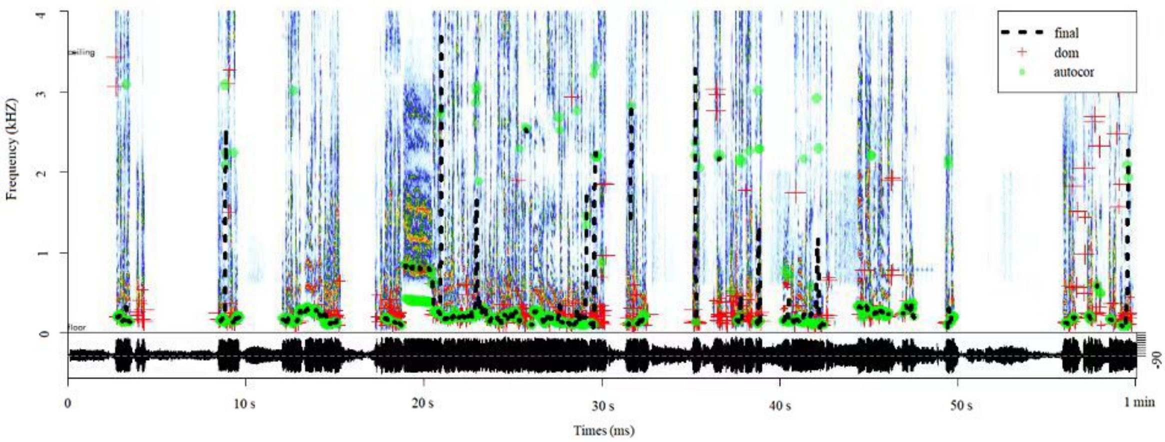
The audio files were converted to WAV format to allow the processing of richer audio data; WAV files have a much higher audio quality than MP3 files and are truer and more accurate to the original sound, although the file sizes are substantially larger. Audio samples were converted from stereo to mono by the single-channel processing of the software used (Anikin, 2019). Voice and music characteristics in the data were split via the Ultimate Vocal Remover using the algorithm from Takahashi and Mitsufuji (2017).

Data processing: audio feature extraction

The tempo (in beats per minute, BPM) of the music (non-voice) audio file for every advertisement was identified via Tunebat (http://tunebat.com/). Tunebat uses a proprietary algorithm for music analysis to detect the BPM of each song. It has a high level of accuracy and is used by many professionals in the music industry, maintaining an extensive database of processed Spotify music files for reference.

Spectrograms of voice and music, such as Super Bowl ads, are illustrated in Figures 3 and 4. The ad is one minute long, as indicated on the Time (x) axis. The samples' frequency (in kHz) is shown in the y-axis, from low to high. The frequency of the voice sample (in Figure 3) has been limited to 4 kHz since all notable features appear below this (the ceiling appears at around 3.4 kHz) for clarity. The top element of each figure shows frequency, while the figure's lower (black) element illustrates amplitude. The spectrogram identifies several key audio features using symbols and colors. The overall detected audio signal is indicated by the colored lines or patches in the upper part of the figure, ranging in loudness from blue to red. The green dots show the pitch calculated through autocorrelation. The red crosses display the lowest dominant frequency. This is the lowest frequency and has the highest power in the audio. The black dotted line gives the final calculated pitch. The appearance of the voice varies significantly throughout the advertisement, with some time periods containing no noticeable voice signal, denoted by the large, rectangular white spaces in the spectrogram across frequencies. As seen by the black dotted lines, voice pitch is elevated at some points in the ad, typically about the visual context and message in the ad.

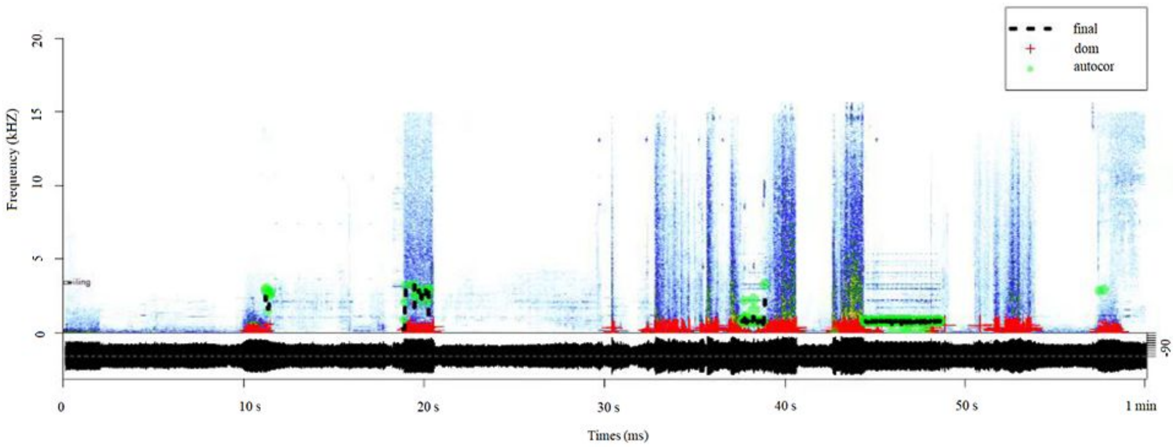
Figure 3. Voice Spectrogram for a Super Bowl Ad



Note: limited to 4 kHz; final = calculated pitch; dom=lowest dominant frequency; autocor=autocorrelation; -90 = -90 dB (amplitude in the lower part of the figure measured in dB).

The spectrogram for the music of the same Super Bowl ad is shown in Figure 4 using the same codes noted above. As seen from the amplitude in the lower part of Figure 4, music occurs throughout the ad, with only relatively small variations in loudness. However, frequency does change at some points in the ad and may be used to emphasize aspects of the voiced or visual advertising message. Typically, music in the ads is selected to underpin the visual and verbal elements of the advertising message being conveyed.

Figure 4. Spectrogram of Music / Accompaniment in a Super Bowl Ad



Note: final = calculated pitch; dom=lowest dominant frequency; autocor=autocorrelation. -90 = -90 dB (amplitude in the lower part of the figure measured in dB).

Merging data sources and data preparation

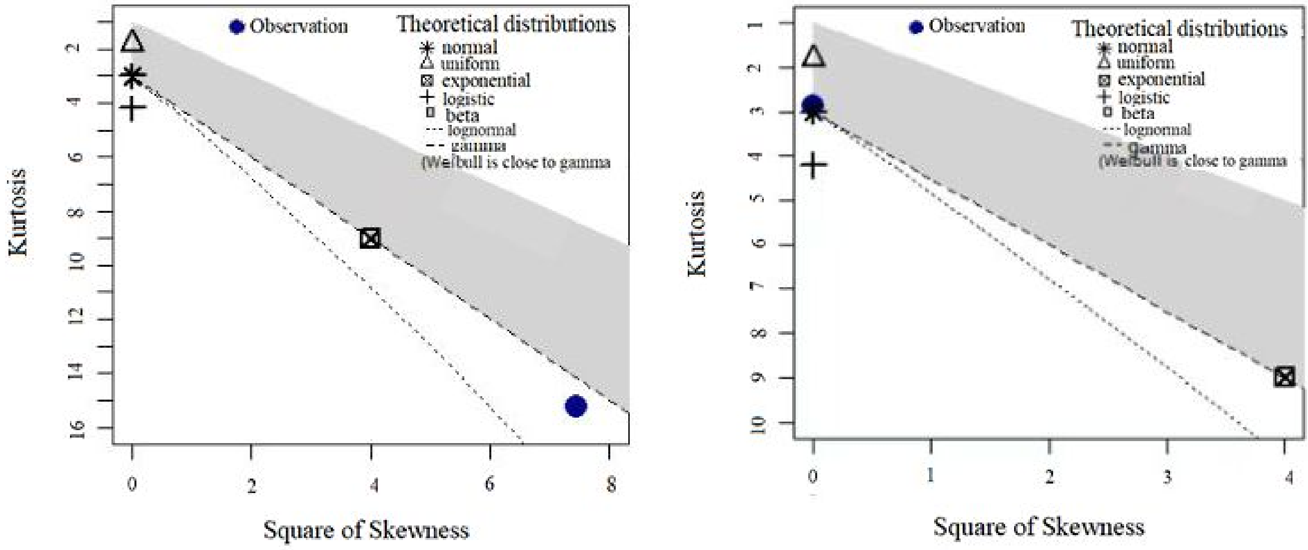
Every Super Bowl ad was coded with a unique identity reference. The unique number was used to merge the different data sources in R. Thus, data on the voice and music features and control variables were consolidated into a single data file.

The liking behavior of an ad (or the net positive effect) was calculated as a function of ad views, as well as the consumer positive (like) and negative (dislike) votes offered for the advertisement. The formula for ad liking behavior i (PLA) is given by:

$$\text{PLA}_\text{i} = \left(\frac{\lambda_i - \vartheta_i}{\tau_i}\right) \qquad (1)$$

where λ_i indicates the total of like votes for video i, ϑ_i refers to the sum of dislike votes for ad i, and τ_i indicates all views of the video advertisement since the first posting. Since the distribution of PLA was found to be heavily positively skewed and leptokurtic (Figure 5a), an Ordered Quartile normalization transformation was applied. The resulting distribution was approximately normal (see Figure 5b). Independent variables were prepared for modeling in the traditional way by scaling and centering them to avoid unit dependence.

Figure 5. Cullen and Frey Graphs.



a. Raw data. **b. Transformed data.**

Descriptive statistics for the overall sample of raw data variables are shown in Table 1, along with the transformed dependent variable. The proportion of voiced frames ranged from 1% to 82%, with an average of 29% of voiced frames in an ad (SD=0.13). The mean of the median intensity of voiced frames was 0.01 (SD=0.02), with a mean standard deviation of 0.03 (SD=0.02). The mean F1 frequency that is typical of male voices was 769.31 Hz (SD=107.77 Hz), while the mean F2 frequency that is more typical of female voices was 1760.76 Hz (SD=190.12 Hz). Regarding music features, the calculated tempo varied considerably from a minimum of 52 BPM to a maximum of 156 BPM, with a mean of 98.2 BPM (SD=27.19 BPM) and a mean tempo-duration interaction of 4699.66 (SD=3027.68). Regarding our control variables, 29% of ad appeals contained a celebrity, 30.4% of the sample included a fear appeal, 27% included a sex appeal, and 69% included a humor appeal. The ad appeal control variables are dichotomous variables. The mean view count for the ads was 1.41 million (SD=11.20 million), with a mean like count of 4146.03 (SD=23,920.40) and a mean dislike count of 833.54 (SD=6948.52).

Table 1. Descriptive Statistics of Sample.

Variables	**n**	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Voice					
*Median Intensity**	241	0.00	0.14	0.01	0.02
*Std. Dev. Intensity**	241	0.00	0.13	0.03	0.02
Mean F1 Frequency (Hz)	241	449.90	1367.87	769.31	107.77
Mean F2 Frequency (Hz)	241	1195.71	2586.29	1760.76	190.12
Proportion of Voice Frames (%)	241	0.01	0.82	0.29	0.13
Music					
Tempo (BPM)	241	52	156	98.20	27.19
Tempo x duration	241	1259.70	26566.95	4699.66	3027.68
Control Variable: Advertising Appeals					
Celebrity Appeal	247	0	1	28.7%	-
Fear Appeal	247	0	1	30.4%	-
Humor Appeal	247	0	1	69.2%	-
Sex Appeal	247	0	1	26.7%	-
Components of Dependent Variable					
View Count	231	10	176373378	1407556.50	11971111.01
Like Count	225	0	275362	4146.03	23920.40
Dislike Count	225	0	92990	833.54	6948.52
Liking Behavior (transformed)	225	-2.84	2.85	0.00	0.99

Note: * Root Mean Square Amplitude per Frame

Model testing and diagnostics

The research model was tested using generalized linear modeling in glmmTMB (Bolker *et al.,* 2009). Models are fitted using maximum likelihood estimation. The fit of the models and terms are compared using the Akaike Information Criterion (AIC), corrected Akaike Information Criterion (AICc), and Log-Likelihood metrics. The residuals of the best-fitting model are further examined using a variety of diagnostics, including the Kolmogorov-Smirnov test to examine the fit of the distribution, a non-parametric dispersion test for over-dispersion, a test for outliers, and a test for heteroskedasticity, all implemented using the DHARMa package (Hartig, 2020).

RESULTS

Table 2 provides the results of testing the research model using generalized linear modeling with a Gaussian distribution and the step-up procedure (Raudenbush and Bryk, 2002), one block of variables at a time. Model 1 includes only the voice variables. Model 2 provides voice and music variables. Finally, model 3 also controls four types of ad appeals. Model 3 is found to have the strongest fit with the data, the lowest Akaike Information Criterion (AIC=593.897) and corrected AIC (AICc=595.664), and the highest log-likelihood (-283.949). Overall, this model explains 17.4% of the variance in the liking behavior of a Super Bowl ad.

Testing Model 3 found that three voice features had a significant positive relationship with the liking behavior of a video advertisement. Median intensity was found to be significantly negatively related to PLA at the 0.1% level ($\beta=-0.32$, $p<0.001$), although the standard deviation of intensity had a non-significant relationship with PLA ($\beta=0.06$, $p=0.403$). This indicates that the data supports H1 and that the average (median) loudness of voice in an advertisement will be negatively related to the liking behavior, demonstrating a dislike for loud voices. Still, that variation in loudness does not affect the liking behavior of the ad, offering no support for H2.

Turning to voice frequencies and the pitch of the speakers, the lower mean F1 frequency range was found to be significantly positively related to the liking behavior of an ad at the 1% level ($\beta=0.25$, $p<0.001$). However, a negative relationship is found; the higher mean F2 frequency range was not significantly related to an ad's liking behavior ($\beta=-0.03$, $p=0.709$). Such finding supports H3a, indicating that lower-pitched voices in an ad tend to be preferred signals for viewers (signal receivers). However, there is no firm evidence that higher-pitched voices are negatively related to the liking behavior of an ad; thus, H3b was not supported.

The proportion of voice frames in each video is significantly positively related to the liking behavior at the 0.1% level (β=0.29, p<0.001). This finding supports H4: an ad's proportion of voice frames will be positively related to liking behavior. More voice stimuli (signals) in an advertisement will create a more significant positive effect on the viewers, the signal receivers.

In terms of music, while the tempo of the music was not found to be significantly related to the liking behavior of an ad (β=-0.05, p=0.436), the interaction of tempo and duration was found to be significantly positively related to PLA at the 5% level (β=0.19, p=0.012). This finding suggests that tempo itself is not an adequate signal in encouraging viewers to like an ad (offering no support for H5a). However, it does support the notion that tempo interacts with the duration of an advertisement, potentially making the ad feel shorter to a consumer; therefore, H5b was supported.

For the control variables, various kinds of ad appeals, the only significant effect on the liking behavior of an ad was that of celebrity appeal (β=0.59, p<0.001). Fear appeals, sex appeals, and humor appeals did not have a significant effect on the liking behavior of an ad.

Table 2. Results of Model Testing.

	Model 1		**Model 2**		**Model 3**	
Predictors	***Estimates (S.E.)***	***p***	***Estimates (S.E.)***	***p***	***Estimates (S.E.)***	***p***
(Intercept)	-0.00 (0.06)	1.000	-0.00 (0.06)	0.983	-0.26 (0.12)	0.034
Voice						
*Median Intensity**	-0.28 (0.10)	0.003	-0.31 (0.09)	0.001	-0.32 (0.09)	<0.001
*Std. Dev. Intensity**	0.07 (0.08)	0.367	0.09 (0.08)	0.268	0.06 (0.08)	0.403

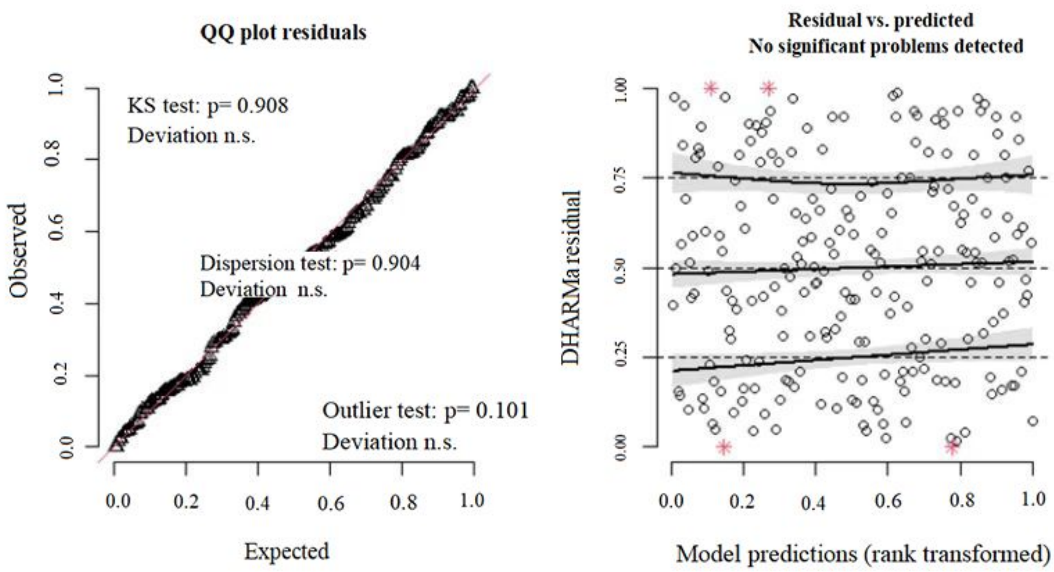
Mean F1 Frequency	0.23 (0.08)	0.002	0.23 (0.07)	0.002	0.25 (0.07)	<0.001
Mean F2 Frequency	-0.04 (0.07)	0.608	-0.01 (0.07)	0.856	-0.03 (0.07)	0.709
Proportion of Voice Frames	0.29 (0.08)	<0.001	0.28 (0.08)	<0.001	0.29 (0.08)	<0.001
Music						
Tempo (BPM)			-0.09 (0.07)	0.176	-0.05 (0.07)	0.436
Tempo x duration			0.26 (0.07)	<0.001	0.19 (0.07)	0.012
Controls: Ad Appeals						
Celebrity Appeal					0.59 (0.14)	<0.001
Fear Appeal					0.22 (0.13)	0.101
Humor Appeal					0.04 (0.15)	0.811
Sex Appeal					-0.01 (0.14)	0.947
Observations	220		220		220	
R^2	0.077		0.124		0.174	
AIC	615.300		604.771		593.897	
AICc	615.828		605.629		595.664	
log-Likelihood	-300.650		-293.386		-283.949	

Model diagnostics.

A comprehensive series of diagnostic tests were performed on the model results using DHARMa, which tests model parameters using advanced simulation and statistics (Hartig, 2020). A summary of tests on the residual diagnostics from Model 3 is shown in Figure 6.

A one-sample Kolmogorov-Smirnov test reveals that the distribution of the model fits the data (D=0.038, p=0.908). A non-parametric dispersion test via the standard deviation of residuals fitted versus simulated found a ratio of observed to simulated data 1.002 (p=0.904), providing no concern over overdispersion. Further, a test for outliers on the data was non-significant, and a heteroskedasticity test for the location of quantiles found no issues (p=0.911). Overall, the model results appear incredibly robust.

Figure 6. Model Diagnostics.



To assess possible multicollinearity problems in the data set, the variance inflation factor (VIF) was calculated and examined (Table 3). The VIF is below 10 in all cases, suggesting multicollinearity is not a problem. The VIF and the 95% confidence intervals for VIF were all less than 3, indicating low multicollinearity between the variables.

Table 3. Test for Multicollinearity: Variance Inflation Factors.

Variable	VIF	VIF 95% CI	Increased S.E.	Tolerance	Tolerance 95% CI
Proportion of Voice Frames	1.63	[1.41, 1.98]	1.28	0.61	[0.51, 0.71]
Mean F1 Frequency	1.46	[1.27, 1.77]	1.21	0.69	[0.57, 0.78]

Mean F2 Frequency	1.43	[1.25, 1.73]	1.20	0.70	[0.58, 0.80]
Median Intensity	2.30	[1.92, 2.82]	1.52	0.44	[0.35, 0.52]
Std. Dev. Intensity	1.68	[1.45, 2.04]	1.30	0.59	[0.49, 0.69]
Tempo (BPM)	1.27	[1.14, 1.55]	1.13	0.79	[0.65, 0.88]
Tempo x duration	1.57	[1.36, 1.90]	1.25	0.64	[0.53, 0.74]
Celebrity Appeal	1.18	[1.07, 1.46]	1.09	0.84	[0.69, 0.93]
Fear Appeal	1.11	[1.03, 1.43]	1.05	0.90	[0.70, 0.97]
Humor Appeal	1.39	[1.22, 1.68]	1.18	0.72	[0.59, 0.82]
Sex Appeal	1.13	[1.04, 1.43]	1.06	0.89	[0.70, 0.96]

DISCUSSION AND CONCLUSIONS

Audio advertising is a vital part of firms' marketing strategies. In terms of voice analytics, every video advertisement consists of numerous features, such as music, voice, and other information, that can be used as signals to consumers to increase the effectiveness of advertising. Brands can bring together different signals to develop their unique profile and identities. Nevertheless, more research is needed to examine the efficacy of the audio features of video embedded in ads using advanced analytical methods. Grounded in signaling theory (Chung and Kalnins, 2001; Spence, 1973), this research affirms the positive effects of low intensity (less loudness), F1 frequency (typical male voices), and the high proportion of voice frames. Furthermore, from the cue congruity lens, the research findings support the interactive effects of tempo and advertisement length, suggesting a higher tempo is a more effective signal for longer advertisements. Consequently, this research offers fruitful theoretical and practical implications.

Theoretical and research implications

This study makes notable contributions to theory and research. First, this research contributes to a new research process for using audio analytics in advertising. The detailed research process outlined can be used for future research examining audio and music from advertisements. Recent advances in analytics offer substantial opportunities for analyzing a greater variety of unstructured data types (Balducci *et al.,* 2018). In particular, the analysis of *actual features of audio* extracted from ads combined with *the actual data on consumer interaction with the videos* offers meaningful insights (Huang and Labroo, 2020; Rodero *et al.,* 2017; Martín-Santana *et al.,* 2015b; Chattopadhyay *et al.,* 2003). While there has been a rapidly increasing amount of research exploring text data, visual elements, and subtle elements (e.g., music) in advertising in recent decades (Dall'Olio and Vakratsas, 2023; Villarroel Ordenes *et al.,* 2019; Ludwig *et al.,* 2013), there is a scarcity of research exploring a more comprehensive variety of modes and media. Responding to Grewal *et al.* (2022) recent call for more multimodal research to explain and predict consumers' behaviors and market developments, this research sheds new light on the audio analytics literature by examining the effects of voice and music signals on the liking behavior of an ad.

Second, our findings provide additional support to the critical role of voice features (e.g., intensity and frequency) as signals in inducing responses from consumers (Biswas *et al.,* 2019; Hagtvedt and Brasel, 2016). Both intensity and frequency features of the actual Super Bowl audio ads will be perceived as signals of emotions that influence consumer interaction with the videos. Echoing the findings of Chattopadhyay *et al.* (2003), our findings highlight the negative relationship between frequency and ad and brand attitudes. Voices with low frequency will be perceived as signals of positive attitudes (Zoghaib, 2017). Voice associations (warm, mature, strong, happy, light, wise, young, and modern) may explain the underlying mechanism of why voice influences brand recall and partly justify customers' attitudes toward the voice (Zoghaib, 2017). Louder voices (with greater sound intensity) tend to be associated with anger, aggression (Juslin and Laukka, 2003), and higher dominance (Tusing and Dillard, 2000). This research demonstrates the importance of verbal cues, particularly frequency and intensity, that customers may use to infer the emotions associated with ads.

Third, the study surfaces a new theoretical association: the interactive effect between tempo and duration on the liking behavior of an ad. The efficacy of music as a signal to consumers is associated with the degree to which the message in an ad is congruent (Kim *et al.,* 2021). While previous research has noted the effect of music tempo on the effectiveness of ads, Martín-Santana *et al.* (2015a) and Zander (2006) proposed that music tempo alone does not show a significant direct effect on an affective level. A faster tempo is associated with longer waiting times (Kim and Zauberman, 2019) and longer duration time (Caldwell and Hibbert, 2002). We extend the advertising literature by demonstrating the interactive effect of music tempo and duration on the liking behavior of an ad. The findings help to explain why fast-tempo and longer ads in real business practice are more effective than other combinations in generating positive attitudes.

Implications for practice

Based on the findings, this research generates practical suggestions for advertisers, especially with respect to sporting events. First, this research confirms the benefits of the voice frames embedded in advertisements (Model 3 found an effect size of 0.29 for the percent of voice frames, significant at the 0.1% level). The greater the percentage of video frames with a human voice, the higher the liking behavior for an ad. The human voice is an effective signal in increasing the liking behavior of an ad and should be used as an essential device in leveraging the value of advertising. Advertising practitioners are encouraged to significantly increase the proportion of human voice frames in their advertisements.

Second, the research findings recognized the vital role of voice features as signals to consumers, as they provide a means for advertisers to create more effective advertisements that customers will like and generally prefer. For example, a quieter voice with a higher F1 frequency has a positive association with the liking behavior of consumers (Model 3 found an effect size of 0.25 for Mean F1 frequency and an effect size of -0.32 for median intensity, both significant at the 0.1% level). Consumers favor less loudness and, most importantly, a high F1 frequency (a typical male vocal tract). These findings are of value to advertisers who may elevate the liking behavior of consumers by carefully selecting spokespersons for their productions that contribute to reducing the level of intensity and using F1 frequencies that are typical of male voices (mean of 769.31 Hz in our sample).

Third, this research confirmed a significant interaction effect between music tempo and duration on liking behavior (Model 3 found an effect size of 0.19, significant at the 5% level). Interestingly, the music tempo itself does not necessarily lead to ad-liking behavior. However, advertisers must be aware that the music tempo should be a clear fit with the length of an ad. When advertisers try to identify the most appropriate music to use, long advertisements fit better with music of a fast tempo.

Fourth, the chosen spokesperson needs to fit with the advertisements but is likely to be a male celebrity with a quieter voice. It is recommended to have at least a third of the advertisement as verbal content, with those with more than half verbal content being most effective (the mean proportion of voice frames is 29%, with a standard deviation of 13%).

Limitations and future research

Despite the meaningful practical and theoretical implications that this research generates, there is still space for future studies to investigate the impacts of audio analytics further. First, this study focused only on Super Bowl advertisements. Future studies may test whether the research findings apply in the context of other sporting events, such as the Olympics and World Cup football. Second, future research may also examine whether the effectiveness of acoustic features of videos differs via the classification of the type of product or service offered in an advertisement.

Third, this study focuses on the frequency and amplitude of voice as critical signals. Therefore, future investigations should examine the effects of additional aspects of soundwaves, such as time (e.g., the average time per word) and spectrum (Sueur, 2018). Using different combinations of acoustic features of voices to represent the brands may induce various perceptions and images (Kim *et al.,* 2021). Such a "brand voiceprint" may influence consumer behavior toward a brand and its products. Thus, future research may explore whether the brand voiceprint is a valuable segmentation criterion for predicting customers' preferences toward various categories of products. Fourth, this research does not consider the congruity effect among multiple elements. For example, we do not account for the fit between the music and human voice features of the spokesperson and various components in advertisements, including colors, the age of the spokesperson, and the characteristics of the target audience. Future research may explore the possible interactions between these various elements in advertising. Finally, celebrity appeal is the only appeal with a significant relationship with ad-liking behavior. Thus, further research could examine additional appeals not covered in this research to identify any that may have been overlooked and investigate aspects of celebrities that affect liking to understand better their interactions with voice and music characteristics in ads.

Acknowledgment

This study is based on the authors' conference paper "Understanding Consumer Advertising via Audio Analytics of Sports Videos," presented at the 2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) authored by Barnes, S. and Wang, W. (2023).

References

Akram, Z., McClelland, A. and Furnham, A. (2018), "The effect of fear-inducing content on memory for advertisements and on retroactive and proactive interference of programme information", *Applied Cognitive Psychology*, Vol. 32 No. 4, pp.413-419.

Anikin, A.S. (2019), "An open-source tool for synthesizing nonverbal vocalizations", *Behavior Research Methods,* Vol. 51*,* pp.2778-2792.

Antonetti, P., Baines, P. and Walker, L. (2015), "From elicitation to consumption: Assessing the longitudinal effectiveness of negative emotional appeals in social marketing", *Journal of Marketing Management*, Vol. 31 No. 9, pp.940–969.

Apple, W., Streeter, L.A. and Krauss, R.M. (1979), "Effects of pitch and speech rate on personal attributions", *Journal of Personality and Social Psychology*, Vol. 37 No. 5, p.715.

Balducci, B. and Marinova, D. (2018), "Unstructured data in marketing", *Journal of the Academy of Marketing Science*, Vol. 46 No. 4, pp.557–590.

Bannister, S. (2020), "A vigilance explanation of musical chills? Effects of loudness and brightness manipulations", *Music & Science*, Vol. 3, pp.1-7.

Belch, G.E. and Belch, M.A. (2013), "A content analysis study of the use of celebrity endorsers in magazine advertising", *International Journal of Advertising,* Vol. 32 No. 3, pp.369-389.

Berger, J. (2019), "Signaling can increase consumers' willingness to pay for green products: Theoretical model and experimental evidence", *Journal of Consumer Behaviour*, Vol. 18 No. 3, pp.233-246.

Biehal, G., Stephens, D. and Curlo, E. (1992), "Attitude toward the ad and brand choice", *Journal of Advertising*, Vol. 21 No. 3, pp.19-36.

Biswas, D., Lund, K. and Szocs, C. (2019), "Sounds like a healthy retail atmospheric strategy: Effects of ambient music and background noise on food sales", *Journal of the Academy of Marketing Science*, Vol. 47 No. 1, pp.37-55.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H. and White, J.S.S. (2009), "Generalized linear mixed models: A practical guide for ecology and evolution", *Trends in Ecology and Evolution*, Vol. 24 No. 3, pp.127-135.

Buttle, H., Raymond, J.E. and Danziger, S. (2000), "Do famous faces capture attention?" *Advances in Consumer Research*, Vol. 27, p.245.

Caldwell, C. and Hibbert, S.A. (2002), "The influence of music tempo and musical preference on restaurant patrons' behavior", *Psychology & Marketing*, Vol. 19 No. 11, pp.895-917.

Chandrasekaran, D., Srinivasan, R. and Sihi, D. (2018), "Effects of offline ad content on online brand search: Insights from Super Bowl advertising", *Journal of the Academy of Marketing Science*, Vol. 46, pp.403-430.

Chattopadhyay, A., Dahl, D.W., Ritchie, R.J.B. and Shahin, K.N. (2003), "Hearing voices: The impact of announcer speech characteristics on consumer response to broadcast advertising", *Journal of Consumer Psychology*, Vol. 13 No. 3, pp.198-204.

Choong, P., Filbeck, G., Tompkins, D.L. and Ashman, T.D. (2003), "An event study approach to evaluating the economic returns of advertising in the Super Bowl", *Academy of Marketing Studies Journal*, Vol. 7 No. 1, pp.89-99.

Chung, W. and Kalnins, A. (2001), "Agglomeration effects and performance: A test of the Texas lodging industry", *Strategic Management Journal*, Vol. 22 No. 10, pp.969-988.

Connelly, B.L., Certo, S.T., Ireland, R.D. and Reutzel, C.R. (2011), "Signaling theory: A review and assessment", *Journal of Management*, Vol. 37 No. 1, pp.39-67.

Dahl, D., Sengupta, J. and Vohs, K. (2009), "Sex in advertising: Gender differences and the role of relationship commitment", *Journal of Consumer Research*, Vol. 36 No. 2, pp.1-31.
Dall'Olio, F. and Vakratsas, D. (2023), "The impact of advertising creative strategy on advertising elasticity", *Journal of Marketing*, Vol. 87 No. 1, pp.26-44.
Dawar, N. (2018), "Marketing in the age of Alexa", *Harvard Business Review*, Vol. 96 No. 3, pp.80-86.
De Pelsmacker, P. and Geuens, M. (1999), "The advertising effectiveness of different levels of intensity of humor and warmth and the moderating role of top of mind awareness and degree of product use", *Journal of Marketing Communications*, Vol. 5 No. 3, pp.113-129.
Eastman, J.K., Iyer, R. and Wiggenhorn, J.M. (2010), "The short-term impact of Super Bowl advertising on stock prices: An exploratory event study", *Journal of Applied Business Research*, Vol. 26 No. 6, pp.69-84.
Eisend, M. (2009), "A meta-analysis of humor in advertising", *Journal of the Academy of Marketing Science*, Vol. 37, pp.191-203.
Fehle, F., Tsyplakov, S. and Zdorovtsov, V. (2005), "Can companies influence investor behaviour through advertising? Super Bowl commercials and stock returns", *European Financial Management*, Vol. 11 No. 5, pp.625-647.
Feng, S., Rajneesh S. and Monique B., (2014), "Does classical music relieve math anxiety? Role of tempo on price computation avoidance", *Psychology & Marketing*, Vol. 31 No. 7, pp.489-499.
Fürst, A., Pečornik, N. and Binder, C. (2021), "All or nothing in sensory marketing: must all or only some sensory attributes be congruent with a product's primary function?", Journal of Retailing, Vol. 97 No. 3, pp.439-458.
Google (2016), "Where people turn to view Super Bowl ads they may have missed", available at: https://www.thinkwithgoogle.com/marketing-strategies/video/super-bowl-advertising-view-statistics/ (accessed 21 February 2022).
Grewal, D., Gotlieb, J. and Marmorstein, H. (1994), "The moderating effects of message framing and source credibility on the price-perceived risk relationship", *Journal of Consumer Research,* Vol. 21 No. 1, pp.145-153.
Grewal, D., Herhausen, D., Ludwig, S. and Ordenes, F.V. (2022), "The future of digital communication research: Considering dynamics and multimodality", *Journal of Retailing*, Vol. 98 No. 2, pp.224-240.
Guha, A., Bressgott, T., Grewal, D., Mahr, D., Wetzels, M. and Schweiger, E. (2023), "How artificiality and intelligence affect voice assistant evaluations", *Journal of the Academy of Marketing Science*, Vol. 51 No. 4, pp.843-866.
Guo, X., Deng, H., Zhang, S. and Chen, G. (2020), "Signals of competence and warmth on e-commerce platforms", *Data and Information Management*, Vol. 4 No. 2, pp.81-93.
Guyer, J.J., Fabrigar, L.R. and Vaughan-Johnston, T.I. (2019), "Speech rate, intonation, and pitch: Investigating the bias and cue effects of vocal confidence on persuasion", *Personality and Social Psychology Bulletin*, Vol. 45 No. 3, pp.389-405.
Hagtvedt, H. and Brasel, S.A. (2016), "Cross-modal communication: Sound frequency influences consumer responses to color lightness", *Journal of Marketing Research*, Vol. 53 No. 4, pp.551-562.
Hartig F. (2020), "DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models", R package version 0.3.2.0., available at: https://CRAN.R-project.org/package=DHARMa (accessed 28 February 2022).

Hendriks, H. and Strick, M. (2019), "A laughing matter? How humor in alcohol ads influences interpersonal communication and persuasion", *Health Communication*, Vol. 35 No. 14, pp.1821-1829.

Hildebrand, C., Efthymiou, F., Busquet, F., Hampton, W.H., Hoffman, D.L. and Novak, T.P. (2020), "Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications", *Journal of Business Research*, Vol. 121, pp.364-374.

Hofacker, C.F., Malthouse, E.C. and Sultan, F. (2016), "Big data and consumer behavior: Imminent opportunities", *Journal of Consumer Marketing*, Vol. 33 No. 2, pp.89-97.

Holbrook, M.B. and Batra, R. (1987), "Assessing the role of emotions as mediators of consumer responses to advertising", *Journal of Consumer Research*, Vol. 14 No. 3, pp.404-420.

Huang, X. and Labroo, A. A. (2020), "Cueing morality: The effect of high-pitched music on healthy choice", *Journal of Marketing*, Vol. 84 No. 6, pp.130-143.

Jain, S., Basu, S., Dwivedi, Y.K. and Kaur, S. (2022), "Interactive voice assistants – does brand credibility assuage privacy risks?" *Journal of Business Research*, Vol. 139, pp.701-717.

Juslin, P.N. and Laukka, P. (2003), "Communication of emotions in vocal expression and music performance: Different channels, same code?" *Psychological Bulletin*, Vol. 129 No. 5, pp.770-814.

Kamiloğlu, R.G., Fischer, A.H. and Sauter, D.A. (2020), "Good vibrations: A review of vocal expressions of positive emotions", *Psychonomic Bulletin & Review*, Vol. 27 No. 2, pp.237-265.

Kanner, B. (2004), The Super Bowl of advertising: How the commercials won the game, UNC Press Books, Chapel Hill.

Kao, K.C., Rao Hill, S. and Troshani, I. (2020), "Effects of cue congruence and perceived cue authenticity in online group buying", *Internet Research*, Vol. 30 No. 3, pp.945-970.

Kellaris, J.J. and Kent, R.J. (1992), "The influence of music on consumers' temporal perceptions: Does time fly when you're having fun?" *Journal of Consumer Psychology*, Vol. 1 No. 4, pp.365-376.

Kim, H.J., Wang, Y. and Ding, M. (2021), "Brand voiceprint", *Customer Needs and Solutions*, Vol. 8, pp.123-136.

Kim, J. and Morris, J. D. (2003), "The effect of advertising on the market value of firms: Empirical evidence from the Super Bowl ads", *Journal of Targeting, Measurement, and Analysis for Marketing*, Vol. 12 No. 1, pp.53-65.

Kim, J.W., Denton, L.T. and Wang, T. (2015), "Assessing stock market response to the release of ad meter rankings of super bowl T.V. commercials", *International Journal of Integrated Marketing Communications*, Vol. 7 No. 1, pp.15-22.

Kim, K. and Zauberman, G. (2019), "The effect of music tempo on consumer impatience in intertemporal decisions", *European Journal of Marketing,* Vol. 53 No. 3, pp.504-523.

Knoeferle, K.M., Knoeferle, P., Velasco, C. and Spence, C. (2016), "Multisensory brand search: How the meaning of sounds guides consumers' visual attention", *Journal of Experimental Psychology: Applied*, Vol. 22 No. 2, p.196.

Ko, S.J., Sadler, M.S. and Galinsky, A.D. (2015), "The sound of power: Conveying and detecting hierarchical rank through voice", *Psychological Science*, Vol. 26 No. 1, pp.3-14.

Krishna, A., Elder, R.S. and Caldara, C. (2010), "Feminine to smell but masculine to touch? Multisensory congruence and its effect on the aesthetic experience", *Journal of Consumer Psychology*, Vol. 20 No. 4, pp.410-418.

Krisjanous, J., Ashill, N.J., Eccarius, K. and Carruthers, J. (2013), “Scared stiff? The effectiveness of threat appeals in counseling services advertising to high-anxiety students”, *Psychology & Marketing*, Vol. 30 No. 10, pp.874-890.
LaTour, M. and Henthorne, T. (1994), “Ethical judgements of sexual appeals in print advertising”, *Journal of Advertising*, Vol. 23 No. 3, pp.81-90.
Lee, J.G. and Ko, K.A. (2021), “The market responses to Super Bowl advertising: The role of product type and multiple executions”, *Sustainability*, Vol. 13 No. 13, p.7127.
Leongómez, J.D., Pisanski, K., Reby, D., Sauter, D., Lavan, N., Perlman, M. and Varella Valentova, J. (2021), “Voice modulation: From origin and mechanism to social impact”, *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 376 No.1840, 20200386.
Li, X., Shi, M. and Wang, X.S. (2019), “Video mining: Measuring visual information using automatic methods”, *International Journal of Research in Marketing*, Vol. 36 No. 2, pp.216-231.
Lortie, C.L., Rivard, J., Thibeault, M. and Tremblay, P. (2017), “The moderating effect of frequent singing on voice aging”, *Journal of Voice*, Vol. 31 No. 1, pp.1-12.
Lowe, M.L. and Haws, K.L. (2017), “Sounds big: The effects of acoustic pitch on product perceptions”, *Journal of Marketing Research*, Vol. 54 No. 2, pp.331-346.
Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E.C., Wetzels, M. and Pfann, G. (2013), “More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates”, *Journal of Marketing*, Vol. 77 No. 1, pp.87-103.
MacKenzie, S.B., Lutz, R.J. and Belch, G.E. (1986), “The role of attitude toward the ad as a mediator of advertising effectiveness: A test of competing explanations”, *Journal of Marketing Research*, Vol. 23 No. 2, pp.130-143.
Martín-Santana, J.D., Muela-Molina, C., Reinares-Lara, E. and Rodríguez-Guerra, M. (2015a), “Effectiveness of radio spokesperson's gender, vocal pitch and accent and the use of music in radio advertising”, *Business Research Quarterly*, Vol. 18 No. 3, pp.143-160.
Martín-Santana, J.D., Reinares-Lara, E. and Muela-Molina, C. (2015b), “Music in radio advertising: Effects on radio spokesperson credibility and advertising effectiveness”, *Psychology of Music*, Vol. 43 No. 6, pp.763-778.
Moraes, C., Kerrigan, F. and McCann, R. (2019), “Positive shock: A consumer ethical judgement perspective”, *Journal of Business Ethics*, Vol. 165, pp.735-751.
Morales, A.C., Wu, E.C. and Fitzsimons, G.J. (2012), “How disgust enhances the effectiveness of fear appeals”, *Journal of Marketing Research*, Vol. 49 No. 3, pp.383-393.
Nail, J. (2007), “Visibility versus surprise: Which drives the greatest discussion of Super Bowl advertisements?”, *Journal of Advertising Research*, Vol. 47 No. 4, pp.412-419.
Newell, S.J. and Henderson, K.V. (1998), “Super Bowl advertising: Field testing the importance of advertisement frequency, length and placement on recall”, *Journal of Marketing Communications*, Vol. 4 No. 4, pp.237-248.
Noh, Y., Kim, H., Kim, K. and Cheong, Y. (2021), “The relationship between consumer engagements with Super Bowl ad tweets and the advertisers’ official Twitter accounts: a panel data analysis of 2019 Super Bowl advertiser”, *International Journal of Advertising*, Vol. 40 No. 6, pp.852-869.
North, A.C., Sheridan, L.P. and Areni, C. S. (2016), “Music congruity effects on product memory, perception, and choice”, *Journal of Retailing*, Vol. 92 No. 1, pp.83-95.
Oakes, S. (2003), “Demographic and sponsorship considerations for jazz and classical music festivals”, *Service Industries Journal*, Vol. 23 No. 3, pp.165-178.

Oakes, S. (2007), "Evaluating empirical research into music in advertising: A congruity perspective", *Journal of Advertising Research*, Vol. 47 No. 1, pp.38-50.
Oakes, S. and North, A.C. (2008), "Using music to influence cognitive and affective responses in queues of low and high crowd density", *Journal of Marketing Management,* Vol. 24 No. 5-6, pp.589-602.
Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K. and Sorokowska, A. (2017), "Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults", *Psychonomic Bulletin & Review*, Vol. 24 No. 3, pp.856-862.
Pantoja, F. and Borges, A. (2021), "Background music tempo effects on food evaluations and purchase intentions", *Journal of Retailing and Consumer Services*, Vol. 63, 102730.
Pathak, A., Calvert, G.A. and Lim, E.A. (2017), "How the linguistic characteristics of a brand name can affect its luxury appeal", *International Journal of Market Research*, Vol. 59 No. 5, pp.567-600.
Poels, K. and Dewitte, S. (2019), "The role of emotions in advertising: A call to action", *Journal of Advertising*, Vol. 48 No. 1, pp.81-90.
Purohit, D. and Srivastava, J. (2001), "Effect of manufacturer reputation, retailer reputation, and product warranty on consumer judgments of product quality: A cue diagnosticity framework", *Journal of Consumer Psychology*, Vol. 10 No. 3, pp.123-134.
Raithel, S., Taylor, C.R. and Hock, S.J. (2016), "Are Super Bowl ads a super waste of money? Examining the intermediary roles of customer-based brand equity and customer equity effects", *Journal of Business Research*, Vol. 69 No. 9, pp.3788-3794.
Raudenbush, S.W. and Bryk, A.S. (2002), *Hierarchical linear models: Applications and data analysis methods*, Second Edition, Thousand Oaks, Sage.
Rázuri, J.G. (2015), "Decision-making content of an agent affected by emotional feedback provided by capture of human's emotions through a bimodal system", *International Journal of Computer Science Issues*, Vol. 12 No. 6, pp.1-20.
Rodero, E., Potter, R.F. and Prieto, P. (2017), "Pitch range variations improve cognitive processing of audio messages", *Human Communication Research*, Vol. 43 No. 3, pp.397-413.
Roozen, I. and Claeys, C. (2010), "The relative effectiveness of celebrity endorsement for print advertisement", *Review of Business and Economic Literature,* Vol. 55 No. 1, pp.76-90.
Scherer, K.R. and Giles, H. (1979), *Social markers in speech*, Cambridge University Press, Paris.
Sengupta, J. and Dahl, D. (2008), "Gender related reactions to gratuitous sex appeals in advertising", *Journal of Consumer Psychology*, Vol. 18 No. 1, pp.62-78.
Severn, J., Belch, G. and Belch, M. (1990), "The effects of sexual and non-sexual advertising appeals and information level on cognitive processing and communication effectiveness", *Journal of Advertising*, Vol. 19 No. 1, pp.14-22.
Shan, M., Zhu, Z., Song, C. and Chen, H.A. (2023), "The effectiveness of advertising appeals: A culturally-derived power perspective", *Journal of Retailing and Consumer Services*, Vol. 75, 103522.
Spangenberg, E.R., Grohmann, B. and Sprott, D.E. (2005), "It's beginning to smell a lot like Christmas: the interactive effects of ambient scent and music in a retail setting", *Journal of Business Research*, Vol. 58 No. 11, pp.1583-1589.
Spence, M. (1973), "Job market signaling", *The Quarterly Journal of Economics*, Vol. 87 No. 3, pp.355-374.
Statista (2021), "*Combined time spent with Super Bowl ads on YouTube on the Game Day from 2016 to 2019 (June 3rd, 2021)*", Statista, available at:

https://www.statista.com/statistics/671858/time-spent-super-bowl-ads-youtube/ (accessed March 3, 2022).

Stewart, D.W. and Punj, G.N. (1998), "Effects of using a nonverbal (musical) cue on recall and playback of television advertising: Implications for advertising tracking", *Journal of Business Research*, Vol. 42 No. 1, pp.39-51.

Sueur, J. (2018), *Sound analysis and synthesis with R*, Springer, Cham.

Sun, Y., Luo, B., Wang, S. and Fang, W. (2020), "What you see is meaningful: Does green advertising change the intentions of consumers to purchase eco-labeled products?", *Business Strategy and the Environment*, Vol. 30 No. 1, pp.694-704.

Sunaga, T. (2018), "How the sound frequency of background music influences consumers' perceptions and decision making", *Psychology & Marketing*, Vol. 35 No. 4, pp.253-267.

Takahashi, N. and Mitsufuji, Y. (2017), "Multi-scale multi-band DenseNets for audio source separation", in *2017 IEEE workshop on applications of signal processing to audio and acoustics* (WASPAA), IEEE, New Paltz, NY, pp.21-25.

Tannenbaum, M.B., Hepler, J., Zimmerman, R.S., Saul, L., Jacobs, S., Wilson, K. and Albarracín, D. (2015), "Appealing to fear: A meta-analysis of fear appeal effectiveness and theories", *Psychological Bulletin*, Vol. 141 No, 6, pp.1178–1204.

Tavassoli, N.T. and Lee, Y.H. (2003), "The differential interaction of auditor and visual advertising elements with Chinese and English", *Journal of Marketing Research*, Vol. 40 No. 4, pp.468-480.

Taylor, C.R. (2016), "Some interesting findings about Super Bowl advertising", *International Journal of Advertising*, Vol. 35 No. 2, pp.167-170.

Tellis, G.J., MacInnis, D.J., Tirunillai, S. and Zhang, Y. (2019), "What drives virality (sharing) of online digital content? The critical role of information, emotion, and brand prominence", *Journal of Marketing*, Vol. 83 No. 4, pp.1-20.

Thomas, L., Shane, S. and Weigelt, K. (1999), "An empirical examination of advertising as a signal of product quality", *Journal of Economic Behavior and Organization*, Vol. 37 No. 4, pp.415-430.

Tigue, C.C., Borak, D.J., O'Connor, J.J., Schandl, C. and Feinberg, D.R. (2012), "Voice pitch influences voting behavior", *Evolution and Human Behavior*, Vol. 33 No. 3, pp.210-216.

Tomkovick, C., Yelkur, R., Rozumalski, D., Hofer, A. and Coulombe, C.J. (2011), "Super Bowl ads linked to firm value enhancement", *Journal of Marketing Development & Competitiveness*, Vol. 5 No. 2, pp.29-43.

Tsui, H.C. (2012), "Advertising, quality, and willingness-to-pay: Experimental examination of signaling theory", *Journal of Economic Psychology*, Vol. 33 No. 6, pp.1193-1203.

Tusing, K.J. and Dillard, J. (2000), "The sounds of dominance. vocal precursors of perceived dominance during interpersonal influence", *Human Communication Research*, Vol. 26 No. 1, pp.148-171.

Van Zant, A.B. and Berger, J. (2020), "How the voice persuades", *Journal of Personality and Social Psychology*, Vol. 118 No. 4, pp.661-682.

Villarroel Ordenes, F., Grewal, D., Ludwig, S., De Ruyter, K., Mahr, D. and Wetzels, M. (2019), "Cutting through content clutter: how speech and image acts drive consumer sharing of social media brand messages", *Journal of Consumer Research*, Vol. 45 No. 5, pp.988-1012.

Vijayaraman, M. and Ramadoss, R. M. (2019), "A study on effect of repeated favorite slow tempo songs on heart rate and systolic B.P. of young adults", *International Journal of Physiology, Nutrition and Physical Education*, Vol. 4 No. 1, pp.180-190.

Wang, S.W. and Scheinbaum, A.C. (2018), “Enhancing brand credibility via celebrity endorsement: Trustworthiness trumps attractiveness and expertise”, *Journal of Advertising Research*, Vol. 58 No. 1, pp.16-32.

Wirtz, J.G., Sparks, J.V. and Zimbres, T.M. (2018), “The effect of exposure to sexual appeals in advertisements on memory, attitude, and purchase intention: a meta-analytic review”, *International Journal of Advertising*, Vol. 37 No. 2, pp.168-198.

Wyllie, J., Carlson, J. and Rosenberger, PJ. (2014), “Examining the influence of different levels of sexual-stimuli intensity by gender on advertising effectiveness”, *Journal of Marketing Management*, Vol. 30 No. 7-8, pp.697-718.

Zander, M.F. (2006), “Musical influences in advertising: How music modifies first impressions of product endorsers and brands”, *Psychology of Music*, Vol. 34 No. 4, pp.465-480.

Zhang, Z. (2016), “Mechanics of human voice production and control”, *Journal of the Acoustical Society of America*, Vol. 140 No. 4, p.2614.

Zoghaib, A. (2017), “The Contribution of a brand spokesperson’s voice to consumer-based brand equity”, *Journal of Product & Brand Management*, Vol. 26 No. 5, pp.492-502.

Zoghaib, A. (2019), “Persuasion of voices: The effects of a speaker’s voice characteristics and gender on consumers’ response”, *Recherche et Applications en Marketing*, Vol. 34 No. 4, pp.83-110.

Zhang, J.Z. and Watson IV, G.F. (2020), “Marketing ecosystem: An outside-in view for sustainable advantage”, *Industrial Marketing Management*, Vol. 88, pp.287-304.

Zhu, R. and Meyers-Levy, J. (2005), “Distinguishing between the meanings of music: When background music affects product perceptions”, *Journal of Marketing Research*, Vol. 42 No. 3, pp.333-345.