Journal of Current Issues & Research in Advertising

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ujci20

Predicting Advertising Persuasiveness: A Decision Tree Method for Understanding Emotional (In)Congruence of Ad Placement on YouTube

Taylor Jing Wen, Ching-Hua Chuan, Jing Yang & Wanhsiu Sunny Tsai

To cite this article: Taylor Jing Wen, Ching-Hua Chuan, Jing Yang & Wanhsiu Sunny Tsai (2022) Predicting Advertising Persuasiveness: A Decision Tree Method for Understanding Emotional (In)Congruence of Ad Placement on YouTube, Journal of Current Issues & Research in Advertising, 43:2, 200-218, DOI: 10.1080/10641734.2021.1963356

To link to this article: https://doi.org/10.1080/10641734.2021.1963356

JOURNAL OF CURRENT ISSUES & RESEARCH IN ADVERTISING
2022, VOL. 43, NO. 2, 200–218
https://doi.org/10.1080/10641734.2021.1963356

Routledge
Taylor & Francis Group

Check for updates

Predicting Advertising Persuasiveness: A Decision Tree Method for Understanding Emotional (In)Congruence of Ad Placement on YouTube

Taylor Jing Wen[a], Ching-Hua Chuan[b], Jing Yang[c], and Wanhsiu Sunny Tsai[b]

[a]School of Journalism and Mass Communications, College of Information and Communications, University of South Carolina, Columbia, South Carolina, USA; [b]School of Communication, University of Miami, Coral Gables, Florida, USA; [c]School of Communication, Loyola University Chicago, Chicago, Illinois, USA

ABSTRACT
By applying the computational method of decision trees, this research identifies the most decisive attributes enhancing ad persuasiveness by examining the contextual effects of emotional (in)congruence on ad placement for music videos on YouTube. Findings of this interdisciplinary research not only evaluated key psychological constructs via a computational approach to predict persuasiveness but also extended the theoretical consideration of contextual (in)congruence into the domain of emotion. Methodologically, this study demonstrates the effectiveness of decision trees in exploratory theory testing. Practically, the predictive results from the decision tree model provide much needed strategic guidance to inform advertising design and evaluation for video-sharing websites.

With the recent big data accessibility, hardware advancement, and algorithm evolvement, computational techniques are engendering a paradigm shift in the advertising landscape (Huh and Malthouse 2020). Through constructing models based on users' purchase behaviors, media consumption history, demographic information, and media contextual tags (Malthouse, Maslowska, and Franks 2018), computational systems are increasingly developed with the aim of more efficiently matching advertisements with those who are interested in the information. For example, contextual advertising that delivers targeted advertisements based on the content each consumer is viewing (Zhang and Katona 2012) can be easily automated.

Studies on contextual advertising conceptualizes medium as a contextual cue that implicitly communicates the message. Scholars have studied such media-context effects on consumers' ad evaluations according to either the assimilation or contrast effects (Dahlén 2005). For example, Belanche and colleagues (2017) reported that comparable and congruent context facilitates the efficacy of online video advertising, while Dahlén et al. (2008) observed that placing ads in thematically incongruent and contrasting media could enhance ad processing and evaluation. It is important to note that context congruence has been narrowly defined in the literature as the degree of similarity between the themes in the advertisement and that of the media program (e.g., Furnham, Bergland, and Gunter 2002). While matching editorial content and advertising message in print and on websites have become a common practice (Kononova et al. 2020; Wojdynski and Bang 2016), thematic congruence is difficult to pair with certain types of media, specifically music videos, one of the most watched contents on the powerful advertising platform of YouTube. According to industry data, 60% of American users aged 35 to

CONTACT Taylor Jing Wen jwen2@mailbox.sc.edu School of Journalism and Mass Communications, College of Information and Communications, University of South Carolina, Room 331, 800 Sumter St., Columbia, SC 29208, USA.

54 years old visit YouTube for music-related content at least once a week (Anderson 2018). Even with the aid of context-aware technology, research has shown that human perception of semantic relatedness can be different from the predictions made by machine-learning algorithms in some specific contexts, such as placing a beer advertisement next to an article about drunk driving (Watts and Adriano 2021). Therefore, this study aims to advance research on contextual advertising by proposing a new approach that emphasizes the role of emotional (in)congruence in ad placement.

Recognizing the importance of emotional experience in music consumption (Juslin and Laukka 2004), this study examines the *emotional* (in)congruence between ads and music videos that underlies the emotional flow of the overall media consumption experience. Existing studies have demonstrated the significant impacts of affective advertising context (i.e., mood and emotions) on advertising effectiveness (e.g., Poels and Dewitte 2019). For example, LaTour and LaTour (2009) showed that positive affect aroused by the media context may positively prime consumers' responses to the advertisements. Shapiro and MacInnis (2002) found that a negative emotion evoked by the media context can promote an analytical processing of the advertising messages. However, there exists scarce empirical evidence regarding whether and how the emotional (in)congruence between the ad and the media context influences consumer response to the ad. Since the effect of emotional (in)congruence is under-theorized in digital ad placement literature, this study adopts an exploratory approach to examine the naturally occurring patterns in the empirical data, inductively test the theoretical relationships among relevant constructs, and robustly verify the applicability of the existing advertising theories in this new context.

Specifically, the current study intends to test important theoretical constructs through the computational method of decision trees. Decision trees are one of the most commonly used machine learning algorithms for inductive inference and predictive tasks including classification and regression (Mitchell 1997). Unlike other machine learning algorithms that are often considered to be black boxes that "do not explain their predictions in a way that humans can understand" (Rudin 2019), decision trees possess a unique advantage as the flowchart-like structure clearly shows the decision-making logic for the prediction results. Moreover, distinct from the conventional approach of verifying hypothesized relationships among a relatively limited set of focal variables, decision trees not only can analyze a much wider range of factors all together to identify the most pivotal factors, but can also illustrate the importance of each factor. The structured model of decision trees ranks the priority of factors as well as "cutoff" points for continuous variables (e.g., if the ad is rated as higher than 3.5 in terms of informativeness). In this way, the results can be directly translated into concrete guidelines for application. Decision trees may thus constitute a powerful data-mining tool for theory testing and building.

Extant studies on digital video advertising effectiveness have revealed several critical factors, including the emotion featured in the video ad (Teixeira, Wedel, and Pieters 2012), the thematic congruence between the ad and its media context (Belanche, Flavián, and Pérez-Rueda 2017), and consumers' message involvement (Huang et al. 2013). Moreover, recent studies on YouTube video advertising have identified the important role of advertising value (Sabuncuoğlu-İnanç, Gökaliler, and Gülay 2020), which is driven by informativeness, entertainment, and irritation. Therefore, this study adopts the decision tree method to analyze the above-mentioned message features and consumer perceptual factors to identify the optimal strategy in driving the persuasiveness of video advertising on YouTube.

This study aims to contribute to the advertising literature in four distinct ways. First, by combining decision trees and experimental design, this interdisciplinary approach expands the growing field of computational research in advertising. Second, the study expands prior research on thematic (in)congruence for ad placement by focusing on emotional (in)congruence based on valence and arousal, which is important for media content (e.g., music videos) that is difficult to pair with ads based on thematic (in)congruence. Third, this research proposed and tested a new algorithm that combines decision trees and statistical tests to explore the theoretical relationships between the key factors conceptualized in advertising information processing literature. This approach illuminates not only the most decisive factors but also demonstrates the interactions among the pivotal factors driving the persuasion effect. In this way, the findings provide much needed theoretical explanations for computational research in advertising. Fourth, by focusing on ad placement on YouTube, the findings provide invaluable strategic guidelines to help advertisers and industry professionals craft effective media strategies for this important advertising platform.

Literature review

Contextual advertising and emotional (in)congruence

As a popular format of emerging online advertising, contextual advertising highlights the importance of media context in which the advertisement is being viewed, such as the editorial or programming materials surrounding and preceding the advertisement (Zhang and Katona 2012). Most existing research on contextual advertising applies the ad-context congruence framework based on the information relevance-accessibility model (Baker and Lutz 2000). On the one hand, some studies suggest that an advertisement that is relevant to the information a consumer is seeking will increase message involvement and, in turn, the likelihood of choosing the advertised brand (Wojdynski and Bang 2016). Studies have also found that contextual relevance of an ad increases message involvement and motivation to process ad information, resulting in more favorable consumer responses to the ad (Chun et al. 2014; Kim, Lee, and Chung 2017). On the other hand, others found that choosing a thematically incongruent media context could make the advertisement more interesting and persuasive because such incongruence adds novelty and interest to the ad, leading to more careful processing of the message (Dahlén et al. 2008). To expand prior research that has primarily examined *thematic* (in)congruence, this study explores *emotional* (in)congruence between the ad and the media context. Regardless of the thematic content, the consistency or inconsistency of emotions triggered by the ad and the media program may have different impacts on ad effectiveness.

The construct of emotions in the present study is characterized by two dimensions: valence and arousal. Valence refers to a positive or negative affective state that constitutes a range from extreme pleasure to extreme displeasure (Cohen, Pham, and Andrade 2008). Arousal indicates the level of stimulation of the affective state (Berger and Milkman 2012). Higher levels of arousal denote the greater activation and intensity of feelings whereas lower arousal indicates deactivation and relaxation. Based on valence and arousal, emotional congruence is accordingly defined as the matching of the two dimensions between emotional responses triggered by the ad and the media program. For example, when both the ad and the music video elicit positive valence and high arousal, this is considered as an emotional congruence condition. When consumers experience the emotional state consistently over the course of the ad and the music video, such congruence may enhance the perceived relevance of the ad due to the affect-priming process (Forgas 1992). In contrast, if the ad and the music video elicit varied responses across valence and/or arousal, such an emotional incongruence condition may alert the consumers and trigger the emotion regulation process (Cohen, Pham, and Andrade 2008), which in turn, leads to more careful processing of the message (Dahlén et al. 2008). The present study uses decision trees to understand the key message factors that predict the effects of different ad placements based on emotional (in)congruence. Therefore, one of the key research questions is to understand how the emotional (in)congruence between the ad and the music video, the ad valence, and the ad arousal, influence the predictors selected by the decision tree model.

> RQ1: How do emotional congruence, ad valence, and ad arousal, influence each predictor selected by the decision tree model?

Psychological processing: Message engagement, flow, and emotional intensity

To provide a holistic understanding of the key factors driving consumer response to ads paired with music videos based on emotional (in)contingency, we examined both consumer psychological factors and ad characteristics. The first set of variables related to the information processing literature identifies how consumers are cognitively engaged with and emotionally involved in online video advertising. Three concepts measuring the psychological processes of advertising messages are included as potential predictors of ad persuasiveness.

Message engagement

Message engagement is "the measure of the contextual relevance in which a brand's messages are framed and presented based on its surrounding context" (Wang 2006, 355). Engagement "occurs when a prospective consumer's mind is turned on to a brand idea enhanced by the surrounding context" (Calder and Malthouse 2008, 2). The numerous social media options across multiple platforms (i.e., desktop, tablets, and mobile phones) means that consumers are now connected to information all the time and have a newfound control over their media experiences. In a recent study of digital video advertising, Bellman et al. (2020) showed that the ads that triggered greater arousal may also lead to higher message involvement. Therefore, advertising engagement is a crucial component in understanding persuasive effectiveness of ad placement focusing on emotional (in)congruence (Calder, Malthouse, and Schaedel 2009).

Flow

Flow is defined as the fully immersed state that people experience when they act with total involvement (Csikszentmihalyi and LeFevre 1989). Individuals become absorbed in their activity in the flow state and their focus narrows to the activity itself. Studies show that flow improves the effectiveness of online video advertising (Yang et al. 2014) and purchase intention (e.g., Luna, Peracchio, and de Juan 2002). As flow is critical for entertainment-oriented online content consumption (Moon, Kim, and Armstrong 2014), this construct may be particularly relevant to the study setting of ad pairing for music videos. The level of emotional congruence may affect the likelihood of consumers enjoying and immersing themselves in the persuasive video ads. Therefore, in addition to message engagement, this research explores the role of flow in the contextual effects of emotional (in)congruence in online video advertising.

Emotional intensity

In addition to the above cognition-oriented mechanisms, this study investigates the effects of emotional intensity as an emotion-oriented mechanism on ad persuasiveness. Emotional intensity (Nabi, Gustafson, and Jensen 2018) refers to the strength of the emotions felt during the entire video viewing experience, including both the ad and the music video. According to the excitation transfer theory, the emotional response of one stimulus is caused by a preceding arousing stimulus (Zillmann 1971). This research examines the role of emotional (in)congruence in ad persuasiveness as consumers are constantly exposed to varied combinations of emotional flows (e.g., from positive/high arousal to negative/low arousal). Such emotional experiences may trigger different levels of intensity that could "color" the evaluations of the ad. In particular, emotional intensity has been found to positively impact the acceptance and persuasiveness of online information exchange (Lee et al. 2013).

Advertising value model: Informativeness, entertainment, and irritation

To understand the key advertising characteristics that may drive consumer evaluation, we adopt the widely studied advertising value model. Advertising value is defined as "a subjective evaluation of the relative worth or utility of advertising to consumers" (Ducoffe 1995, 1). Advertising that lacks value tends to result in negative consumer responses, such as 'tuning out' or negative outcomes such as counter-arguing, and ad skipping. In contrast, advertising of high value is likely to contribute to the formation of positive consumer attitudes. The ad value model is based on the Uses and Gratification theory which identifies three perceptual factors affecting ad value: informativeness, entertainment, and irritation (Ducoffe 1995). On the one hand, research shows that perceived advertising value is a key determinant of advertising effectiveness in digital advertising (Xu, Oh, and Teo 2009), such as website advertising, email advertising, and mobile advertising (Cheng et al. 2009). Specifically, when consumers find that the advertising message is more informative, more entertaining, and less irritating, they will perceive the ad as having higher value (Liu et al. 2012). On the other hand, when assessing the validity of the ad value model for the social media advertising, Logan et al. (2012) demonstrated that consumers perceived ad entertainment higher for social media ads but ad informativeness higher for television ads. They also found that ad irritation directly influenced attitude toward advertising but did not play a significant role in ad value assessment. The present study reexamined the Ducoffe's ad value model in digital video advertising using decision tree to explore the inherent patterns of the data and the relationships between these dimensions.

Informativeness refers to the ability to provide relevant information (Oh and Xu 2003). Consumers who value advertising information may consider information provision as one of the need-satisfying functions derived from media communication (Blanco, Blasco, and Azorín 2010). Shavitt and colleagues (1998) found that consumers regarded information as a positive aspect of advertising when they learned about new products, specific product benefits, and comparative product information. By contrast, the value of entertainment lies in its ability to fulfill the needs for escapism, diversion, esthetic enjoyment, and/or emotional release (Ducoffe 1995). Therefore, entertainment is a positive source of advertising value and is crucial to the effectiveness of online advertising (Ducoffe 1996). A high degree of pleasure and involvement during interaction with computer-based media leads to concurrent subjective perceptions of positive affect (Hoffman and Novak 1996). Further, enjoyment associated with advertisements plays the most critical role in accounting for overall ad attitudes (Shavitt, Lowrey, and Haefner 1998). While the informativeness and entertainment variables are positive antecedents of advertising value (Ducoffe 1995, 1996), the irritation variable serves as a negative predictor. Ad irritation leads to a general reduction in perceived ad value and effectiveness (Brehm 1966). Consumers' irritation with advertising is related to the content of advertising, the sheer amount of advertising clutter, deception and privacy concerns (Greyser 1973).

Another perceptual variable pertinent to ad value is visual attractiveness that reflects the degree to which a person believes an ad is esthetically pleasing to the eye (Van der Heijden 2003). Attractive visual elements in the ad can capture consumer attention while driving a desirable consumption experience (Schmid 1998). Consumers may derive esthetic values from ads (Bashir et al. 2018). Therefore, the visual attractiveness may also contribute to ad persuasiveness.

Intrusiveness

Particularly relevant to the study focus of YouTube advertising that has often been criticized as intrusive, this study pays special attention to the perceived ad intrusiveness, which is the cognitive assessment of the degree to which an ad interferes with individuals' cognitive processes and hinders their objectives (Edwards, Li, and Lee 2002). Perceived intrusiveness gauges the level of distraction that an ad causes, especially for online video viewing experiences. An ad perceived as intrusive may evoke annoyance and prompt negative emotional responses toward the ad, resulting in ad avoidance (Rejón-Guardia and Martínez-López 2014). Prior research that examines ways to reduce intrusiveness of online advertising indicates that increasing relevance might be a useful strategy for advertisers (Edwards, Li, and Lee 2002). This is because ads that are self-relevant may be regarded as more interesting and less intrusive to consumers' online experience. Emotional (in)congruence could potentially serve as a cue to signal the relevancy of or spark the audience's curiosity about the ad content, which in turn may mitigate ad intrusiveness and facilitate persuasiveness. Given the sundry factors analyzed in this study, we propose the following research question:

> RQ2: What are the significant predictors of the persuasiveness of video advertising selected by the decision tree model?

Method

Experiment design and stimuli selection

To examine the research questions proposed above, a 2 (congruent vs. incongruent) x 2 (positive valence vs. negative valence) x 2 (high arousal vs. low arousal) between-subject online experiment was conducted. This project is a part of a broad, ongoing effort to explore the incorporation of computational methods in advertising research to understand the media-contextual effects of emotional (in)congruence. In a previous study, we used supervised machine learning methods to predict the valence and arousal elicited by video ads. The stimuli pool included 48 30-seconds video ads collected from YouTube. These ads were selected because of the dominant role music plays in the overall appeal. The topics of these ads included commercial goods (e.g., Coca Cola and Samsung Phone), services (e.g., Progressive Insurance and Letgo App), and public service advertisements (e.g., climate change and child abuse). To pair the ads with music videos, we obtained the music videos from the DEAP dataset (Koelstra et al. 2011), which is the most widely used dataset in affective computing. Specifically, the multimodal algorithm of the DEAP dataset analyzes both visual features (e.g., colorfulness and lighting) and audio features (e.g., rhythm, harmony, and timbre), and connects the patterns in these features with the viewers' affective score to predict the valence and arousal scores of the music videos. The algorithm trained with the DEAP dataset was then used to predict the valence and arousal scores of each of the 48 ads.

With the valence and arousal scores for the music videos reported in the DEAP dataset, four pairs of music videos and video ads based on emotional congruence and another 4 pairs for emotional incongruence were generated. For example, we paired a positive and high-arousal ad with a positive and high-arousal music video that has the similar valence and arousal scores from the DEAP dataset. They were thus selected as the congruence pair for the positive valence and high arousal condition. For the incongruence conditions, the ads are paired with the music videos with the opposite scores on the valence and arousal dimensions. For example, the positive and high-arousal ad was paired with the negative and low-arousal music video as one of the incongruence conditions.

Since the algorithm was trained using the music videos from the DEAP dataset, a manipulation check test ($N = 118$) was conducted to validate the accuracy of the affective score predictions for the 4 selected video ads. The two-way ANOVA tests showed that video ad groups had strong main effects on self-report valence and arousal scores. Individuals reported higher valence scores when exposed to the algorithm-predicted high-valence ads ($M = 6.25$, SD = 1.97) than the low-valence ads ($M = 3.34$, SD = 1.79), $F(1, 114) = 71.413$, $p < .001$, $\eta_p^2 = .378$. Similarly, the self-report arousal score was higher when individuals were exposed to the algorithm-predicted high-arousal ads ($M = 6.05$, SD = 1.79) than the low-arousal ads ($M = 3.86$, SD = 1.75), $F(1, 114) = 42.252$, $p < .001$, $\eta_p^2 = 0.276$. Specifically, (a) the positive and high arousal ad induced positive valence ($M = 6.40$) and higher arousal ($M = 6.20$); (b) the positive and low arousal ad induced positive valence ($M = 6.10$) and lower arousal ($M = 4.24$); (c) the negative and high arousal ad induced negative valence ($M = 3.79$) and higher arousal ($M = 5.90$); and (d) the negative and low arousal ad induced negative valence ($M = 2.90$) and lower arousal ($M = 3.50$). No other effects were significant, all *ps* > .05, suggesting that the four video ads manipulated valence and arousal largely independently. The algorithm prediction was considered accurate.

Participants and procedure

We recruited 714 participants (63% males) from Amazon Mechanical Turk. The majority of the participants were non-Hispanic White (65.8%) and had at least a bachelor's degree (63.4%). Their ages ranged from 19 to 92 ($M = 36.33$, SD = 10.29). After giving consent to participate in the study, participants first answered screening questions about whether they have watched videos on video sharing sites like YouTube. Participants were then randomly assigned to one of the eight experimental conditions. We set a timer for each condition to make sure that the participants watched the entire stimulus, including both the music video and the video ad to avoid invalid or fake responses. After viewing the entire video, participants were directed to an online questionnaire to provide their perceptions and evaluations of the ad.

Measures

To comprehensively evaluate the influence of the manipulated variables, various psychological factors, and media contextual factors on ad persuasiveness, the decision tree predictive model analyzes all these factors as predictors together at the same time. Table 1 lists all the predictors included in the decision tree analyses.

Decision trees and random forests

Decision trees have been recognized as effective tools for data mining. Decision trees are particularly useful for extracting the importance of attributes or factors, as exemplified by De Oña, Eboli, and Mazzulla (2014) public transportation study that analyzed survey data to identify the importance of frequency, speed, and comfort, respectively, in predicting customer satisfaction. Specifically, the predictive model is built by partitioning the predictor space (constructed with all predictors as dimensions) into a set of regions or nodes, which are sought to be

Table 1. All variables considered in the experiment.

Variable types Manipulated Variables	Variable names	Measurements/Citation	Cronbach's α
	Perceived Congruence	Becker-Olsen, Cudmore, and Hill (2006)	.977
	Valence	Predicted by machine learning algorithms Measures defined in Koelstra et al. (2011)	
	Arousal	Predicted by machine learning algorithms Measures defined in Koelstra et al. (2011)	
Target Variable	Persuasiveness	Kees, Burton, and Tangari (2010)	.944
Potential Predictors	Flow	Kim and Han (2014)	.885
	Emotional Intensity	Nabi, Gustafson, and Jensen (2018)	.887
	Message Engagement	Wang (2006)	.916
	Ad value	Ducoffe (1995)	.960
	Informativeness	Ducoffe (1995)	.948
	Entertainment	Ducoffe (1995)	.966
	Irritation	Li, Edwards, and Lee (2002)	.969
	Attractiveness	Verhagen et al. (2012)	.961
	Intrusiveness	Li, Edwards, and Lee (2002)	.947
	Familiarity	Machleit, Allen, and Madden (1993)	.888
	Involvement	Zaichkowsky (1985)	.909
	Social Media Dependency	Tsai and Men (2013)	.855
	Music Preferences	Spears and Singh (2004)	.978
Demographics	Age	Stated in years	
	Gender	Males/Females/Others	

homogenous with respect to the prediction outcome. Given a prediction target (i.e., the dependent variable), the decision tree algorithm iteratively divides the dataset into partitions by selecting an attribute as the splitter and assigns a value to the data points in each partition such that the prediction errors are minimized. In each iteration, all the attributes are examined and the one that best improves the prediction performance is selected as the splitter to divide a partition into two or more homogeneous subregions. The prediction performance can be measured as data impurity (i.e., heterogeneity in each partition using Gini index) or information gain using entropy (Raileanu and Stoffel 2004). This study adopted the Classification and Regression Tree (CART) algorithm (Steinberg and Colla 2009), one of the most prominent tree algorithms (Kern, Klausch, and Kreuter 2019). To predict the perceived persuasiveness as a numeric value, CART is used for a regression task with the splitting rule that minimizes the mean squared error for each partition.

In addition to finding the splitters and splitting criteria, CART also reports an importance score for attributes. The importance score indicates how useful an attribute is in building the tree, and it is the total reduction of the error brought by the attribute, normalized or weighted by the number of data points in the partitions. For example, the total error reduced by attribute x_1 can be calculated as follows:

$$\frac{1}{\left(N_2+N_3+N_4\right)}\sum_{i\in\left(N_2\cup N_3\cup N_4\right)}\left(y_i-\frac{1}{\left(N_2+N_3+N_4\right)}\sum_{i\in\left(N_2\cup N_3\cup N_4\right)}y_i\right)^2-$$
$$\left[\frac{1}{N_2}\sum_{i\in N_2}\left(y_i-\overline{y_2}\right)^2+\frac{1}{\left(N_3+N_4\right)}\sum_{i\in\left(N_3\cup N_4\right)}\left(y_i-\frac{1}{\left(N_3+N_4\right)}\sum_{i\in\left(N_3\cup N_4\right)}y_i\right)^2\right]$$

Although a prediction model can be built by one single decision tree, such an approach is vulnerable to small changes in the training data. The hierarchical nature of the tree growing process means that a change in one split point affects the remaining splits down the tree (Kern, Klausch, and Kreuter 2019). This limitation is often addressed by building multiple trees to make the prediction in an ensemble manner to improve the generalizability of the model. This approach is called random forests (Breiman 2001), which randomly draws bootstrap samples from the training dataset with replacements (a technique called bagging) to build trees in the forest. The final prediction is aggregated from the result of each tree's prediction via a statistical function such as voting or averaging.

The proposed algorithm: Combining decision trees with statistical tests

In order to meaningfully present the relations between the various predictors and the target value of perceived ad persuasiveness on the decision tree, we developed an algorithm by combining decision trees with statistical tests. Decision trees, like most machine learning methods, focus only on the prediction accuracy but not on identifying relationships between the variables. The proposed algorithm differs from traditional decision trees because it first uses a random forest to select predictors that statistically significantly improve the results before building the decision tree. The algorithm also conducts Ordinary Least Squares (OLS) multiple linear regression on the selected splitter and the manipulated variables in the resulting partition of samples in order to show how a particular predictor is influenced by the experimental setting. The algorithm is summarized in Table 2.

In this study, the input parameters were set to $n = 30$, $k = 10$, and $p = 0.05$. The stop criteria (stop_criteria in Table 2) were used to stop the growing of the decision tree. The criteria can be set using the pruning strategy based on the improvement of the model (e.g., the minimum reduction of errors), or based on the tree's structural parameters such as the maximum depth of the tree and the minimum number of data points in the terminal node.

Table 2. The proposed algorithm with decision trees and statistical tests.

Input data: manipulated variables demographics, other predictors, and the predicted target value.
Input parameters: number of trees in the random forest, number of folds for cross validation, and p-value threshold for statistical tests.

// Using random forest and paired t-test to select predictors

$$X \leftarrow VUDUP$$

Build a random forest with trees with attributes in X for the target value in T. Sort the attributes $x_1, x_2, \ldots x_{|x|}$ in Xin a descending order based on the attributes' importance () score such that $s(x_1) \geq s(x_2) \geq \ldots \geq s(x_{|x|})$ where $|X|$ represents the total number of attributes in X.

$$S \leftarrow \{x_1\}$$

for $i \leftarrow 2$ to $|X|$ **do**

$E_0 \leftarrow$ The mean-square-error for *k*-fold cross validation results with S as attributes for the *n*-tree random forest
$E_1 \leftarrow$ The mean-square-error for *k*-fold cross validation results with S as attributes for the *n*-tree random forest
$p_i \leftarrow$ the p-value of the paired t-test on E_0 and E_1

if $p_i \leq p$ **then**

$S \leftarrow SUx_i$

end if
end for
// Building the decision tree with the selected predictors using CART and OLS
while stop_criteria == false **do**
for every terminal node in the tree **do**
Find the splitter x_i from and the splitting value with CART algorithm
With the samples in the terminal node, conduct OLS regression on x_i with variables in V,D and $p - x_i$
end for
end while

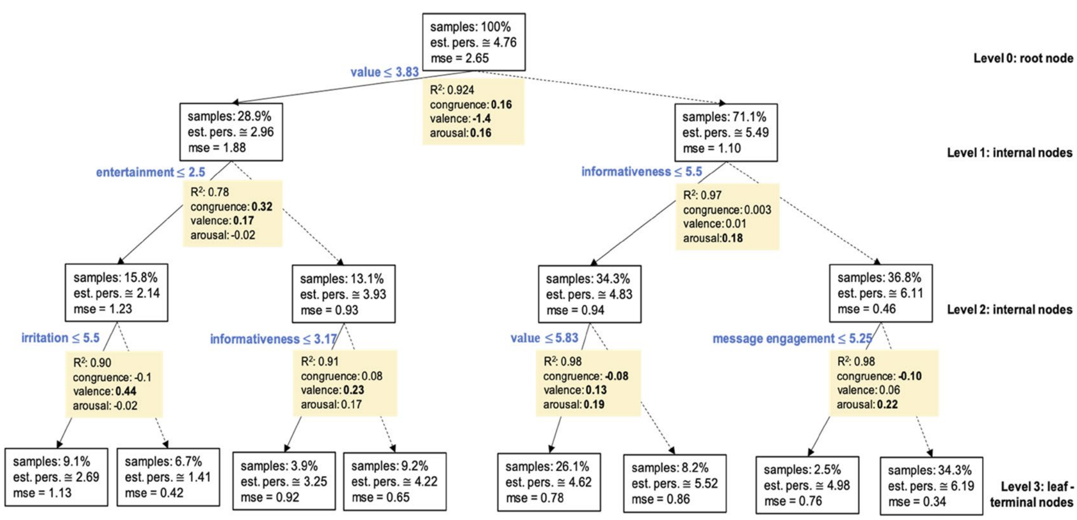

Figure 1. The constructed decision tree for predicting persuasiveness in the dataset.

Interpretations of the model

Figure 1 shows the decision tree built based on the selected attributes in this study. Each transparent rectangle represents a node in the tree, displaying information, including the size of the partition (a percentage showing the number of data points in the partition against the total number of data points), estimated persuasiveness value (est. pers.), and mean square error (mse) based on the estimated persuasiveness value. The equation next to the rectangle shows

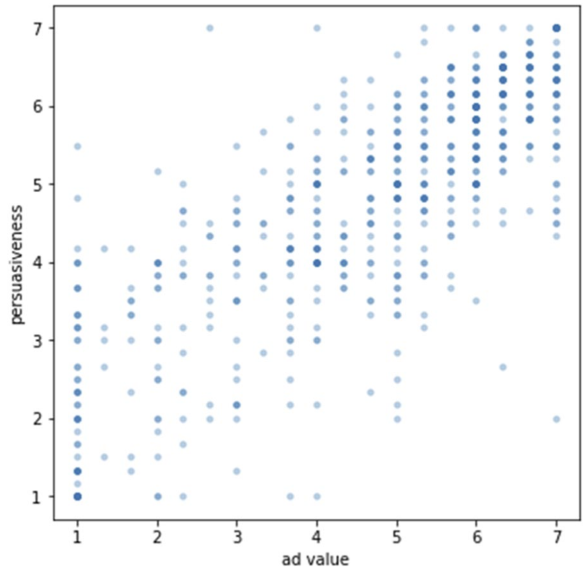

Figure 2. Data distributions of ad value (the root node) and persuasiveness.

the attribute that the algorithm selects as the splitter. For example, as shown in Figure 1, the first attribute selected by the algorithm is the perceived ad value: the algorithm splits the entire dataset into two partitions - one consisting of data points with perceived ad value less than or equal to 3.83 (pointed to by the solid arrow) and the other with ad value greater than 3.83 (pointed to by the dashed arrow). To show how the splitter attribute is impacted by the manipulated variables, the OLS multiple regression is conducted as shown in the colored rectangle below the node, with R^2 value and coefficients for the manipulated variables. Coefficients highlighted in bold have *p*-value less than 0.05.

To examine the manner in which the algorithm splits the dataset, for example at the root node of the tree (level 0), Figure 2 presents a scatter plot showing all the data with two attributes: ad value versus persuasiveness. Note that the darker points in Figure 2 present multiple participants' ratings as they share the same values in these two attributes. A quick glance at the figure can confirm why the algorithm splits the data at the threshold of 3.83 in ad value as the majority of the data points with ad value higher than the threshold have high persuasiveness score while points with ad value lower than the threshold report low persuasiveness. In addition, it can be seen that more darker points appear in the region above the threshold. This explains the percentage of the samples (71.1% versus 28.9%) at the nodes in level 1 as more data points are above 3.83 in ad value in Figure 2.

To understand why the algorithm selects ad value at the root node, Figure 3 shows the scatter plots for other attributes at the same level (level 0). Attributes are placed in a decreasing order in the figure based on the importance score calculated by the algorithm. It can be observed that the importance score informs how the value of an attribute can be used to determine persuasiveness. For example, similar to ad value, informativeness in (a) is roughly correlated with persuasiveness but the pattern is not as clear as in Figure 2. The rest of the attributes in Figure 3 do not show any direct relationship with persuasiveness.

After the algorithm splits the dataset into two groups based on the ad value threshold of 3.83 at level 0, the algorithm examines the data in each group and selects another attribute to further improve the modeling of persuasiveness. Figure 4 shows the data at the two nodes in level 1: (a) data points having ad value less than or equal to 3.83 and (b) data points having ad value greater than 3.83. For data points with smaller ad value scores, entertainment is found to be the most effective attribute to determine persuasiveness, with a threshold of 2.5 annotated as a vertical line in Figure 4(a). For data points with higher ad value scores, informativeness is

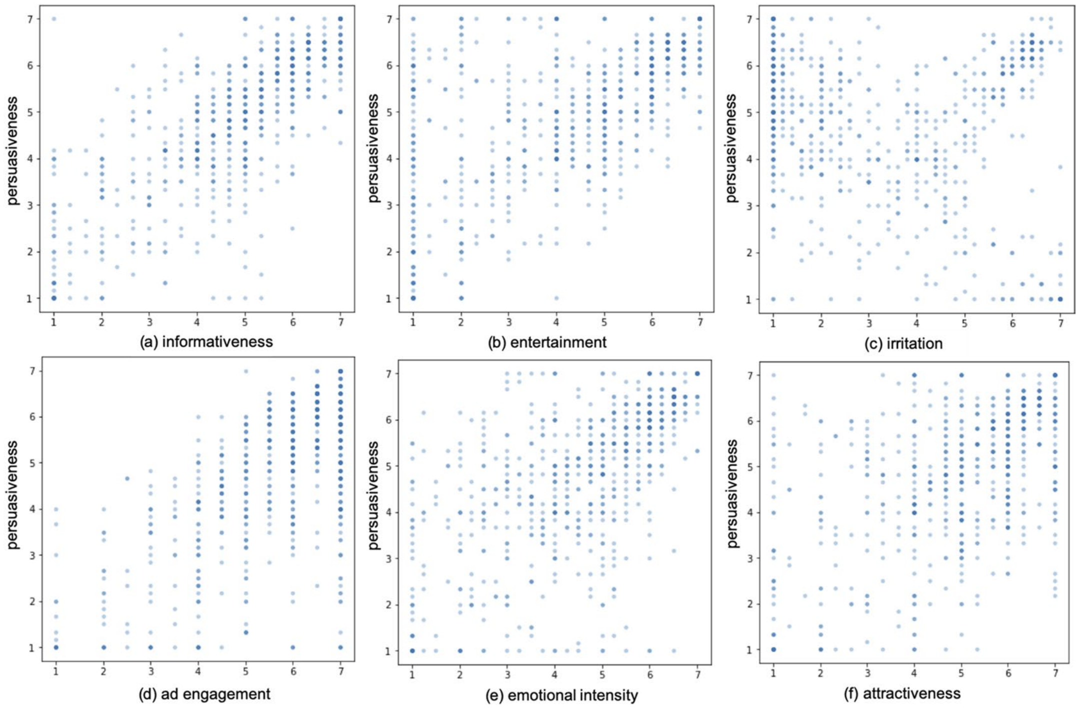

Figure 3. Data distributions of all predicted attributes at level 0.

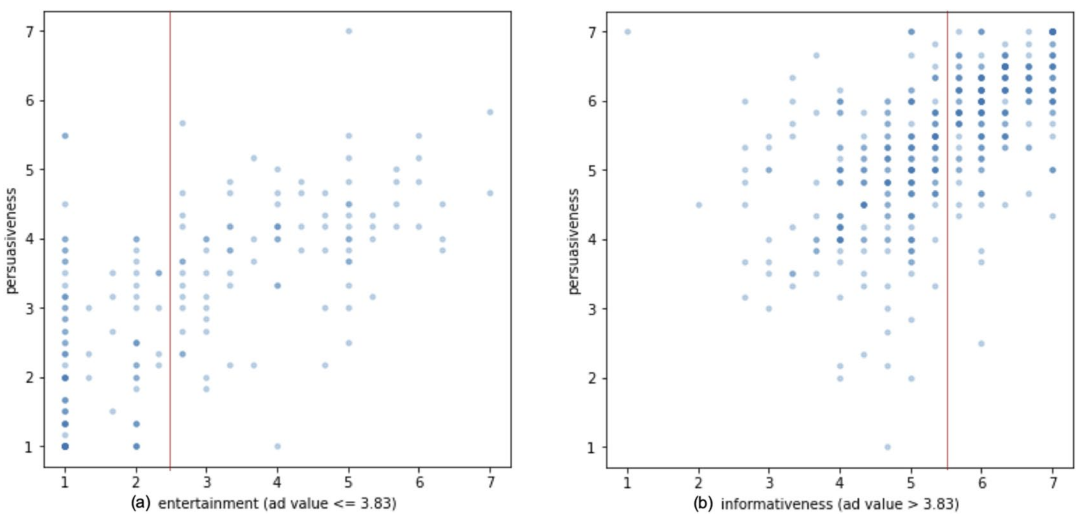

Figure 4. Cutoff points of two nodes (entertainment and informativeness) at level 1.

the identified indicator for persuasiveness. Note that the algorithm improves the modeling accuracy of persuasiveness at each split. For example, at level 1 as shown in Figure 4(b) before the split using informativeness, the estimated persuasiveness of all data points is 5.49, resulting mse of 1.10. After splitting on the threshold of 5.5 on informativeness, the persuasiveness of the subgroup higher than the threshold is estimated as 6.11 (mse = 0.46) and the persuasiveness of the lower subgroup is estimated as 4.83 (mse = 0.94). For each subgroup, the algorithm continues to split them into smaller groups by finding the most effective attribute based on the data distribution in the group, and ultimately, creates the decision tree as shown in Figure 1.

Results

Research question 1 (R1) asked how the manipulated independent variables influenced the predictors selected by the decision tree model. Research question 2 (R2) asked which specific variables would be chosen by the decision tree model as determinant factors influencing the persuasiveness of video advertising. Based on the structural results of the decision trees, we would first discuss the findings of R2 to identify the key predictors of each level in the decision tree and then further discuss how our manipulated variables influenced the chosen predictors of each level as proposed by R1.

Based on results of the decision tree predictions, perceived ad value is the most decisive factor that influences perceived persuasiveness of video advertising. This is the root node of the decision tree (level 0). Then, the decision tree model splits the data by the ad value score of 3.83 into two partitions, which can be further divided into sub-partitions by using different predictors at internal nodes. The different branches of the decision tree leading to varied prediction outcomes, from right to left, are summarized as four sets of rules as shown in Table 3. Therefore, to answer R2, ad value, ad entertainment, ad informativeness, ad irritation, and message engagement were the determinant factors that influenced the persuasiveness of video advertising.

Specific to the study focus of emotional (in)congruence and to answer RQ1, the OLS regression is included in the algorithm to demonstrate the significant effects of the manipulated variables on the important factors leading to different branches of the decision tree. At the root node (level 0), emotional congruence and level of arousal positively influence the perceived ad value whereas valence negatively influences ad value. That is, when the video ad and the media program elicit similarly negative and high arousal emotions, consumers are more likely to perceive a higher ad value. At level 1, on the right branch of the root node, arousal has a positive effect on perceived informativeness. By contrast, on the left branch of the root node at level 1, congruence and valence have positive effects on ad entertainment such that the more congruent the ad and the media program are on the positive emotional dimension, the greater the entertainment value. At level 2, on the right branch under the internal node with informativeness as the splitter, message engagement is negatively influenced by congruence but positively influenced by arousal. The left branch under informativeness, ad value, emerged again as the decisive factor, which is negatively influenced by congruence and positively influenced by valence and arousal. In other words, consumers may find the ad that triggers positive and high arousal emotions while being placed in an incongruent media program (negative and low arousal) possesses higher ad value, even though the ad has relatively low perceived informativeness. At level 2, on the right branch under entertainment, ad valence positively influences perceived ad

Table 3. The four sets of rules derived from the decision tree model.

if perceived ad value > 3.83, and
-- if perceived informativeness of ad > 5.5, and
---- if message engagement > 5.25 ⇒ perceived persuasiveness is estimated as 6.19 **[highest]**
---- else (message engagement ≤ 5.25) ⇒ perceived persuasiveness is estimated as 4.98

if perceived ad value > 3.83, and
-- if perceived informativeness of ad ≤ 5.5, and
---- if perceived ad value > 5.83 ⇒ perceived persuasiveness is estimated as 5.52 **[2nd highest]**
---- else (perceived ad value ≤ 5.83) ⇒ perceived persuasiveness is estimated as 4.62

if perceived ad value ≤ 3.83, and
-- if entertainment > 2.5, and
---- if perceived informativeness of ad > 3.17 ⇒ perceived persuasiveness is estimated as 4.22
---- else (perceived informativeness of ad ≤ 3.17) ⇒ perceived persuasiveness is estimated as 3.25

if perceived ad value ≤ 3.83, and
-- if entertainment ≤ 2.5, and
---- if ad irritation > 5.5 ⇒ perceived persuasiveness is estimated as 1.41 **[lowest]**
---- else (ad irritation ≤ 5.5) ⇒ perceived persuasiveness is estimated as 2.69 **[2nd lowest]**

informativeness. On the left branch under entertainment, ad valence also positively influences perceived ad irritation. In other words, when the entertainment value is high, positive emotion can enhance informativeness. However, when the entertainment value is low, positive emotion can increase irritation.

Discussion and conclusion

Theoretical implications

As pointed out by Huh and Malthouse (2020), most of the literature on computational advertising are published in CS/MIS journals and conference proceedings, but little work has been found in advertising literature. Little connection has been made between the psychological factors developed through robust theory testing in advertising and the algorithms developed in computing. Consequently, most existing computational advertising systems explore the "what" in optimizing advertising effectiveness, but fall short in answering the "why" and "how" of the system's overall decision-making process. To exemplify the use of computational methods for theory building and testing, the current study begins with a broad conceptual framework incorporating key psychological factors identified in the advertising literature into our computational analysis. Our approach thus enhances and highlights the explainability of algorithms. In addition, our findings can potentially combat the threat of loss/inaccuracy of user data, and provide theoretical insights to identify the most crucial factors emerging from the data.

The present research identified the most decisive attributes enhancing consumer evaluation of ad persuasiveness by using the computational method of decision trees to examine the contextual effects of emotional (in)congruence for ad placement with music videos on YouTube. The biggest advantage of the decision tree method lies in its ability to test multiple models and concepts together and identify the crucial attributes which emerge from the internal patterns of the data. Compared to the conventional deductive approach of testing hypothesized relationships among several key variables, this study's decision tree predictive model is more inclusive, which allows for inductive prediction for theory verification. More importantly, the decision tree model ranks the priority of the attributes and detects the cutoff points to illustrate the distinctive interactions among the attributes in the dataset.

Among all the attributes analyzed, we found ad value to be the most influential attribute appearing at the root node. That is, the likelihood that consumers will be persuaded depends on how much value they perceive in the ad. Ad value has been demonstrated as a fundamental attribute to predict ad effectiveness in varied research contexts, such as location-based advertising (Xu, Oh, and Teo 2009) and mobile advertisement (Liu et al. 2012). The current study provides new evidence for the significance of ad value to online video advertising, specifically when investigating the contextual influences of emotional (in)congruence. Moreover, the three dimensions (i.e., informativeness, entertainment, and irritation) in the ad value model were extracted as additional influential attributes that appear at the lower levels below ad value. Hence, the decision tree illuminates the varied conditions in which ad value interacts with these three crucial attributes to predict persuasiveness. It is important to note that although no predetermined relationships were hypothesized among variables measured in this study, the computational decision tree model selected all the variables in the ad value model based on the patterns picked up from the data. Therefore, the results provide invaluable empirical evidence that confirms the intercorrelations between variables in the ad value model. Our findings also suggested that future research regarding online video advertising should consider integrating the ad value model as key constructs or including these variables as covariates.

Moreover, message engagement emerged as another important predictor to persuasiveness. Specifically, if an ad is perceived as high in ad value, informativeness, and engagement, this further increases its persuasiveness. Based on uses and gratification theory, Kim et al. (2015) observed the positive correlations between informativeness and entertainment and advertising engagement. Research shows that consumers are more likely to engage with an ad and brand with messages of higher value, and consumers tend to engage with a brand to receive information-oriented value (Wang and Calder 2009). Complementing prior research, our results suggest the beneficial role of message engagement in enhancing ad persuasiveness. If message engagement is not sufficiently high (greater than 5.25), the persuasiveness of the message will be significantly decreased. In other words, the positive effects of ad value and informativeness will be less likely to translate into persuasive outcomes if consumers do not fully engage themselves with the message. Future research on online video advertising should take an integrative approach that considers both the perceptual message characteristics and consumers' cognitive involvement with the message.

In addition, the current research contributes to the advertising literature by extending the theoretical consideration of contextual (in)congruence into the domain of emotions. The notion of congruence in digital advertising literature has been narrowly defined from the thematic perspective, that is, the degree of similarity between the themes in the ad and that of the media program (Furnham, Bergland, and Gunter 2002). Beyond the congruence of thematic content that requires cognitive processing, our results demonstrate that the match of emotions elicited by both the media context and the ad can also achieve perceived congruence. From the lens of emotional processing, this study extends the conceptualization of context (in)congruence by emphasizing the under-researched affective components.

Finally, the study findings identified how emotional (in)congruence, valence, and arousal influence ad value, informativeness, entertainment, irritation, and message engagement. Previous research has suggested that ad value could be influenced by media context (Ducoffe 1995, 1996). Consumers who select a particular media vehicle may regard advertising that fits closely with the media editorial environment of that vehicle to be of greater value because it addresses their particular interests (Aaker and Brown 1972). Our results suggest that such a fit may derive from emotional congruence between the media vehicle and the ad. Specifically, when consumers find the media program and the ad are emotionally congruent on the dimension of negative valence and high arousal, perceived ad value tends to be higher. The ad stimuli used in such condition in our study was a PSA about child abuse. It is likely that the negatively arousing emotions elicited by the ad carried over to and are amplified by the emotionally congruent media program (i.e., the music video), which facilitates consumers' perceived value of the ad. Moreover, perceived entertainment is found to be higher when the media context and the ad are congruent on positive valence. In other words, regardless of the arousal dimension, as long as both the media program and the ad trigger positive emotions, consumers will find the ad entertaining. This confirms the affect priming model, suggesting that affect renders congruent information accessible and thus encourages affect-congruent judgments (Forgas 1992). The affect priming model indicates that information with the same valence can be linked together and easily activated by the corresponding affect. In this case, the positive emotions elicited by both the media program and the ad would jointly prime substantially positive responses (Bower and Forgas 2000). Finally, the higher the arousal level triggered by the ad, the more likely consumers are to report higher message engagement and ad informativeness. Arousal relates to the activation and intensity of the felt emotions (Cohen, Pham, and Andrade 2008). The use of arousing emotions in the ad may grab consumers' attention and motivate them to continue focusing on the ad content by emotional stimulation (Berger and Milkman 2012). Similarly, when consumers focus on the ad content itself, they are more likely to appreciate its informational value (Kim et al. 2015).

Methodological implications

One significant contribution of this study is the application of computational methods (i.e., decision trees and random forest) in testing advertising theory (i.e., the inclusion of psychological factors rather than merely behavioral variables). As one of the most commonly used machine learning algorithms for inductive inference and predictive tasks, the method of decision tree has unique advantages in providing a flow-chart like constructed model to show the decision-making logic for the prediction outcome. Unlike conventional hypothesis testing processes which are limited to a set of focal variables, decision trees can analyze a larger number of variables and select the ones that are most influential to the prediction variable. Such a distinction can provide more opportunities for scholars to investigate which theoretical construct, under what condition, plays a more significant role than the others in driving advertising effectiveness. Moreover, this study provided detailed steps for implementing the computational method in testing the psychological factors suggested in advertising theory. It is important to note that although a single tree can determine the important factors from a given dataset, a very deep tree is prone to overfitting, diminishing the generalizability of the predictive model. Although pruning techniques (i.e., cutting back a large, unrestricted tree) can avoid overfitting, prediction accuracy is often compromised. Therefore, instead of building only one tree for prediction, we combined many trees into a robust ensemble via random forest. Random forests select the important factors such that the prediction accuracy is maintained, and model generalizability is improved. Additionally, to ensure the selected factors generate significant improvement in prediction, paired t-test is included in the algorithm. To take advantage of the decision trees' simple and explainable representations, the prediction model is displayed in one decision tree.

Practical implications

One potential threat to the current computational advertising techniques is the heavy reliance on consumers' behavioral data which could be hugely impacted if policy change occurs to restrict usage of consumers' personal data. There is a high possibility that programmatic advertising will no longer be able to target individual consumers based on their behavioral data, and the accuracy and efficiency of these computational advertising systems will decrease if data accessibility is restricted. Therefore, the present study provides advertising professionals with actionable insights of contextual targeting, which focuses on the contextual elements rather than the behavioral data.

The predictive model of decision trees enables advertisers to predict the persuasiveness of their messages based on the rating of the decisive factors before running the ad. Additionally, the decisive factors identified by the decision tree can serve as concrete objectives for creative strategy and design that can be verified in pretests to ensure the effectiveness of the message. In particular, the perceived value of a video advertisement emerged as the most decisive factor driving the persuasiveness of video ads on YouTube. Therefore, advertisers should pay special attention to ensure that the video ad is perceived as at least moderately useful, important, and valuable (greater than 3.83) by the target consumers. To achieve high persuasiveness, the video ad also needs to be highly informative (scoring higher than 5.5 in informativeness) by providing timely and relevant product information and be highly engaging (scoring higher than 5.25 in engagement) by holding consumers' attention and keeping them involved. In other words, ad value, informativeness, and engagement should be the prioritized evaluation objectives during pretests. Specific to ad placement for music videos on YouTube, a video ad that is emotionally congruent with the music video, of high arousal, but negative in valence tends to enhance the important driver of perceived ad value. However, when the perceived ad value is low, the decision tree results also pinpoint additional factors and mechanisms that can boost ad persuasiveness. For example, when designing ads for parity products, which may be difficult to portray as important and relevant to achieve high ad value, advertisers should focus on emotional congruence and valence to boost the entertainment value of the ad, by designing the ad as entertaining, enjoyable, and pleasing (i.e., achieving higher than 2.5 in entertainment value) and then focus on ad valence to make it informative (i.e., achieving higher than 3.17 in informativeness) in order to enhance ad persuasiveness. In other words, based on the structural results of decision trees, strategic guidelines can be developed to inform advertising design and evaluation under different conditions.

Limitations and future studies

This study presents one of the earliest efforts of applying the computational method of decision trees and random forest for advertising theory testing. Several limitations should be considered when interpreting the study results and addressed in future research. The analysis was conducted with data obtained through an online experiment, limiting the generalizability of the study findings. Additionally, while decision trees are effective tools for demonstrating the interactions among the factors and the joint effects of pivotal factors on the prediction outcome, decision trees cannot test causal relationships. Given their advantage of being able to adapt to complex relationships while at the same time being effective in terms of pre-processing effort needs, decision trees have been advocated for survey research (Kern, Klausch, and Kreuter 2019). In particular, because the structure of the decision trees is picked up solely from the data, not specified in advance (i.e., based on theory or prior literature), decision trees can be a powerful tool for understanding emerging, novel, under-theorized phenomena and trends. Additionally, as a powerful data exploration tool, decision trees can help identify key variables and their interactions that can be further verified with statistical techniques. Advertising studies exploring complex or under-researched relationships thus should include decision trees to identify the optimal set of factors to enhance theory development and testing.

ORCID

Taylor Jing Wen http://orcid.org/0000-0002-6967-1617
Ching-Hua Chuan http://orcid.org/0000-0001-5350-6134
Jing Yang http://orcid.org/0000-0001-6393-5012
Wanhsiu Sunny Tsai http://orcid.org/0000-0003-3124-6882

References

Aaker, David A., and Phillip K. Brown. 1972. "Evaluating Vehicle Source Effects." *Journal of Advertising Research 12* (4): 11–16.

Anderson, Matt. 2018. "40 Years After Video Killed the Radio Star, YouTube Is Transforming Music: Here's What Brands Need to Know," *Think with Google*, June. https://www.thinkwithgoogle.com/music-videos/

Baker, William E., and Richard J. Lutz. 2000. "An Empirical Test of an Updated Relevance-Accessibility Model of Advertising Effectiveness." *Journal of Advertising 29* (1): 1–14. doi:10.1080/00913367.2000.10673599.

Bashir, Aqsa, Jing Wen, Eunice Kim, and Jon D. Morris. 2018. "The Role of Consumer Affect on Visual Social Networking Sites: How Consumers Build Brand Relationships." *Journal of Current Issues & Research in Advertising 39* (2): 178–191. doi:10.1080/10641734.2018.1428250.

Becker-Olsen, Karen, Andrew Cudmore, and Ronald Paul Hill. 2006. "The Impact of Perceived Corporate Social Responsibility on Consumer Behavior." *Journal of Business Research 59* (1): 46–53. doi:10.1016/j.jbusres.2005.01.001.

Belanche, Daniel, Carlos Flavián, and Alfredo Pérez-Rueda. 2017. "Understanding Interactive Online Advertising: Congruence and Product Involvement in Highly and Lowly Arousing, Skippable Video Ads." *Journal of Interactive Marketing 37*: 75–88. doi:10.1016/j.intmar.2016.06.004.

Bellman, Steven, Robert F. Potter, Jennifer A. Robinson, and Duane Varan. 2021. "The Effectiveness of Various Video Ad-Choice Formats." *Journal of Marketing Communications 27* (6): 631–650. doi:10.1080/13527266.2020.1753091.

Berger, Jonah, and Katherine L. Milkman. 2012. "What Makes Online Content Viral?" *Journal of Marketing Research 49* (2): 192–205. doi:10.1509/jmr.10.0353.

Blanco, Carlos Flavián, Miguel Guinalíu Blasco, and Isabel Iguacel Azorín. 2010. "Entertainment and Informativeness as Precursory Factors of Successful Mobile Advertising Messages." *Communications of the IBIMA 10*: 1–11. doi:10.5171/2010.130147.

Bower, Gordon H., and Joseph P. Forgas. 2000. "Affect, Memory, and Social Cognition." In *Cognition and Emotion*, edited by E. Eich, J. F. Kihlstrom, G. H. Bower, J. P. Forgas, and P. M. Niedenthal, 87–168. New York: Oxford University Press.

Brehm, JackW. 1966. *A Theory of Psychological Reactance*. New York: Academic Press.
Breiman, Leo. 2001. "Random Forests." *Machine Learning 45* (1): 5–32. doi:10.1023/A:1010933404324.
Calder, Bobby J., and Edward C. Malthouse. 2008. "Media Engagement and Advertising Effectiveness." *In Kellogg on Advertising and Media*, edited by B. J. Calder. Hoboken, 1–37. NJ: Wiley & Sons.
Calder, Bobby J., Edward C. Malthouse, and Ute Schaedel. 2009. "An Experimental Study of the Relationship between Online Engagement and Advertising Effectiveness." *Journal of Interactive Marketing 23* (4): 321–331. doi:10.1016/j.intmar.2009.07.002.
Cheng, Julian Ming-Sung, Charles Blankson, Edward Shih-Tse Wang, and Lily Shui-Lien Chen. 2009. "Consumer Attitudes and Interactive Digital Advertising." *International Journal of Advertising 28* (3): 501–525. doi:10.2501/S0265048709200710.
Chun, Kwang, Yeun, Ji Hee Song, Candice R. Hollenbeck, and Jong-Ho Lee. 2014. "Are Contextual Advertisements Effective? The Moderating Role of Complexity in Banner Advertising." *International Journal of Advertising 33* (2): 351–371. doi:10.2501/IJA-33-2-351-371.
Cohen, Joel B., Michel Tuan Pham, and Eduardo B. Andrade. 2008. "The Nature and Role of Affect in Consumer Behavior." In *Handbook of Consumer Psychology*, edited by Curtis Haugtvedt, Paul Herr, and Frank Kardes, 297–348. Mahwah, NJ: Lawrence Erlbaum.
Csikszentmihalyi, Mihaly, and Judith LeFevre. 1989. "Optimal Experience in Work and Leisure." *Journal of Personality and Social Psychology 56* (5): 815–822. doi:10.1037/0022-3514.56.5.815.
Dahlén, Micael. 2005. "The Medium as a Contextual Cue: Effects of Creative Media Choice." *Journal of Advertising 34* (3): 89–98. doi:10.1080/00913367.2005.10639197.
Dahlén, Micael, Sara Rosengren, Fredrik Törn, and Niclas Öhman. 2008. "Could Placing Ads Wrong Be Right? Advertising Effects of Thematic Incongruence." *Journal of Advertising 37* (3): 57–67. doi:10.2753/JOA0091-3367370305.
De Oña, Rocío, Laura Eboli, and Gabriella Mazzulla. 2014. "Key Factors Affecting Rail Service Quality in the Northern Italy: A Decision Tree Approach." *Transport 29* (1): 75–83. doi:10.3846/16484142.2014.898216.
Ducoffe, Robert H. 1995. "How Consumers Assess the Value of Advertising." *Journal of Current Issues & Research in Advertising 17* (1): 1–18. doi:10.1080/10641734.1995.10505022.
Ducoffe, Robert H. 1996. "Advertising Value and Advertising on the Web." *Journal of Advertising Research 36* (5): 21–32.
Edwards, Steven M., Hairong Li, and Joo-Hyun Lee. 2002. "Forced Exposure and Psychological Reactance: Antecedents and Consequences of the Perceived Intrusiveness of Pop-up Ads." *Journal of Advertising 31* (3): 83–95. doi:10.1080/00913367.2002.10673678.
Forgas, Joseph P. 1992. "Mood and the Perception of Unusual People: Affective Asymmetry in Memory and Social Judgments." *European Journal of Social Psychology 22* (6): 531–547. doi:10.1002/ejsp.2420220603.
Furnham, Adrian, Jenny Bergland, and Barrie Gunter. 2002. "Memory for Television Advertisements as a Function of Advertisement–Programme Congruity." *Applied Cognitive Psychology 16* (5): 525–545. doi:10.1002/acp.812.
Greyser, Stephen A. 1973. "Irritation in Advertising." *Journal of Advertising Research 13*: 3–10.
Hoffman, Donna L., and Thomas P. Novak. 1996. "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations." *Journal of Marketing 60* (3): 50–68. doi:10.2307/1251841.
Huang, Jinsong, Song Su, Liuning Zhou, and Xi Liu. 2013. "Attitude toward the Viral Ad: Expanding Traditional Advertising Models to Interactive Advertising." *Journal of Interactive Marketing 27* (1): 36–46. doi:10.1016/j.intmar.2012.06.001.
Huh, Jisu, and Edward C. Malthouse. 2020. "Advancing Computational Advertising: Conceptualization of the Field and Future Directions." *Journal of Advertising 49* (4): 367–376. doi:10.1080/00913367.2020.1795759.
Juslin, Patrik N., and Petri Laukka. 2004. "Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening." *Journal of New Music Research 33* (3): 217–238. doi:10.1080/0929821042000317813.
Kees, Jeremy, Scot Burton, and Andrea Heintz Tangari. 2010. "The Impact of Regulatory Focus, Temporal Orientation, and Fit on Consumer Responses to Health-Related Advertising." *Journal of Advertising 39* (1): 19–34. doi:10.2753/JOA0091-3367390102.
Kern, Christoph, Thomas Klausch, and Frauke Kreuter. 2019. "Tree-Based Machine Learning Methods for Survey Research." *Survey Research Methods 13* (1): 73–93.
Kim, Jihye, Jaejin Lee, and Yoo Jin Chung. 2017. "Product Type and Spokespersons in Native Advertising–the Role of Congruency and Acceptance." *Journal of Interactive Advertising 17* (2): 109–123. doi:10.1080/15252019.2017.1399838.
Kim, Jooyoung, Jungwon Lee, Samsup Jo, Jaemin Jung, and Jaewon Kang. 2015. "Magazine Reading Experience and Advertising Engagement: A Uses and Gratifications Perspective." *Journalism & Mass Communication Quarterly 92* (1): 179–198.
Kim, Yoo Jung, and Jin Young Han. 2014. "Why Smartphone Advertising Attracts Customers: A Model of Web Advertising, Flow, and Personalization." *Computers in Human Behavior 33*: 256–269. doi:10.1016/j.chb.2014.01.015.
Koelstra, Sander, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras. 2011. "DEAP: A Database for Emotion Analysis Using Physiological Signals." *IEEE Transactions on Affective Computing 3* (1): 18–31. doi:10.1109/T-AFFC.2011.15.

Kononova, Anastasia, Wonkyung Kim, Eunsin Joo, and Kristen Lynch. 2020. “Click, Click, Ad: The Proportion of Relevant (vs. Irrelevant) Ads Matters When Advertising within Paginated Online Content.” *International Journal of Advertising 39* (7): 1031–1058. doi:10.1080/02650487.2020.1732114.

LaTour, Kathryn A., and Michael S. LaTour. 2009. “Positive Mood and Susceptibility to False Advertising.” *Journal of Advertising 38* (3): 127–142. doi:10.2753/JOA0091-3367380309.

Lee, Kun Chang, Kwon, Ohbyung, Choong-Ryuhn Kim, and Gimun Kim. 2013. “Factors Affecting the Intensity of Emotional Expressions in Mobile Communications.” *Online Information Review 37* (1): 114–131. doi:10.1108/14684521311311667.

Li, Hairong, Steven Edwards, and Joo-Hyun Lee. 2002. “Measuring the Intrusiveness of Advertisements: Scale Development and Validation.” *Journal of Advertising 31* (2): 37–47. doi:10.1080/00913367.2002.10673665.

Liu, Chia-Ling, Rudolf R. Sinkovics, Noemi Pezderka, and Parissa Haghirian. 2012. “Determinants of Consumer Perceptions toward Mobile Advertising—A Comparison between Japan and Austria.” *Journal of Interactive Marketing 26* (1): 21–32. doi:10.1016/j.intmar.2011.07.002.

Logan, Kelty, Laura F. Bright, and Harsha Gangadharbatla. 2012. “Facebook versus Television: Advertising Value Perceptions among Females.” *Journal of Research in Interactive Marketing 6* (3): 164–179. doi:10.1108/17505931211274651.

Luna, D., L. A. Peracchio, and M. D. de Juan. 2002. “Cross-Cultural and Cognitive Aspects of Website Navigation.” *Journal of the Academy of Marketing Science 30* (4): 397–410. doi:10.1177/009207002236913.

Machleit, Karen, Chris Allen, and Thomas Madden. 1993. “The Mature Brand and Brand Interest: An Alternative Consequence of Ad-Evoked Affect.” *Journal of Marketing 57* (4): 72–82. doi:10.2307/1252220.

Malthouse, Edward C., Ewa Maslowska, and Judy U. Franks. 2018. “Understanding Programmatic TV Advertising.” *International Journal of Advertising 37* (5): 769–784. doi:10.1080/02650487.2018.1461733.

Mitchell, Tom M. 1997. *Machine Learning*. McGraw-Hill, NY: New York.

Moon, Yun Ji, Woo Gon Kim, and Deborah J. Armstrong. 2014. “Exploring Neuroticism and Extraversion in Flow and User Generated Content Consumption.” *Information & Management 51* (3): 347–358. doi:10.1016/j.im.2014.02.004.

Nabi, Robin, Abel Gustafson, and Risa Jensen. 2018. “Framing Climate Change: Exploring the Role of Emotion in Generating Advocacy Behavior.” *Science Communication 40* (4): 442–468. doi:10.1177/1075547018776019.

Oh, Lih-Bin, and Heng Xu. 2003. “Effects of Multimedia on Mobile Consumer Behavior: An Empirical Study of Location-Aware Advertising.” *ICIS 2003 Proceedings*. 56. https://aisel.aisnet.org/icis2003/56

Poels, Karolien, and Siegfried Dewitte. 2019. “The Role of Emotions in Advertising: A Call to Action.” *Journal of Advertising 48* (1): 81–90. doi:10.1080/00913367.2019.1579688.

Raileanu, Laura, and Kilian Stoffel. 2004. “Theoretical Comparison between the Gini Index and Information Gain Criteria.” *Annals of Mathematics and Artificial Intelligence 41* (1): 77–93. doi:10.1023/B:AMAI.0000018580.96245.c6.

Rejón-Guardia, Francisco, and Francisco J. Martínez-López. 2014. “Online Advertising Intrusiveness and Consumers’ Avoidance Behaviors.” In *Handbook of Strategic E-Business Management*, edited by F. J. Martinez-Lopez, 565–586. Berlin: Springer.

Rudin, Cynthia. 2019. “Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models instead.” *Nature Machine Intelligence 1* (5): 206–215. doi:10.1038/s42256-019-0048-x.

Sabuncuoğlu-İnanç, Ayda, Ebru Gökaliler, and Göker Gülay. 2020. “Do Bumper Ads Bump Consumers?: An Empirical Research on YouTube Video Viewers.” *El Profesional de la Información 29* (1): 1–14. doi:10.3145/epi.2020.ene.14.

Schmid, Jack. 1998. “The 8 Commandments of Design.” *Catalog Age 15* (1): 57–59.

Shapiro, Stewart, and Deborah J. MacInnis. 2002. “Understanding Program-Induced Mood Effects: Decoupling Arousal from Valence.” *Journal of Advertising 31* (4): 15–26. doi:10.1080/00913367.2002.10673682.

Shavitt, Sharon, Pamela Lowrey, and James Haefner. 1998. “Public Attitudes toward Advertising: More Favorable than You Might Think.” *Journal of Advertising Research 38* (4): 7–22.

Spears, Nancy, and Surendra Singh. 2004. “Measuring Attitude toward the Brand and Purchase Intentions.” *Journal of Current Issues & Research in Advertising 26* (2): 53–66. doi:10.1080/10641734.2004.10505164.

Steinberg, Dan, and Phillip Colla. 2009. “CART: Classification and Regression Trees.” *The Top Ten Algorithms in Data Mining 9*: 179.

Teixeira, Thales, Michel Wedel, and Rik Pieters. 2012. “Emotion-Induced Engagement in Internet Video Advertisements.” *Journal of Marketing Research 49* (2): 144–159. doi:10.1509/jmr.10.0207.

Tsai, Wan-Hsiu Sunny, and Linjuan Rita Men. 2013. “Motivations and Antecedents of Consumer Engagement with Brand Pages on Social Networking Sites.” *Journal of Interactive Advertising 13* (2): 76–87. doi:10.1080/15252019.2013.826549.

Van der Heijden, Hans. 2003. “Factors Influencing the Usage of Websites: The Case of a Generic Portal in The Netherlands.” *Information & Management 40* (6): 541–549. doi:10.1016/S0378-7206(02)00079-4.

Verhagen, Tibert, Frans Feldberg, Bart van den Hooff, Selmar Meents, and Jani Merikivi. 2012. “Understanding Users’ Motivations to Engage in Virtual Worlds: A Multipurpose Model and Empirical Testing.” *Computers in Human Behavior 28* (2): 484–495. doi:10.1016/j.chb.2011.10.020.

Wang, Alex. 2006. “Advertising Engagement: A Driver of Message Involvement on Message Effects.” *Journal of Advertising Research 46* (4): 355–368. doi:10.2501/S0021849906060429.

Wang, Jing, and Bobby J. Calder. 2009. “Media Engagement and Advertising: Transportation, Matching, Transference and Intrusion.” *Journal of Consumer Psychology 19* (3): 546–555. doi:10.1016/j.jcps.2009.05.005.
Watts, Jameson, and Anastasia Adriano. 2021. “Uncovering the Sources of Machine-learning Mistakes in Advertising: Contextual Bias in the Evaluation of Semantic Relatedness.” *Journal of Advertising 50* (1): 26–38. doi:10.1080/00913367.2020.1821411.
Wojdynski, Bartosz W., and Hyejin Bang. 2016. “Distraction Effects of Contextual Advertising on Online News Processing: An Eye-Tracking Study.” *Behaviour & Information Technology 35* (8): 654–664. doi:10.1080/0144929X.2016.1177115.
Xu, Heng, Lih-Bin Oh, and Hock-Hai Teo. 2009. “Perceived Effectiveness of Text vs. Multimedia Location-Based Advertising Messaging.” *International Journal of Mobile Communications 7* (2): 154–177. doi:10.1504/IJMC.2009.022440.
Yang, Keng-Chieh, Conna Yang, Chia-Hui Huang, Po-Hong Shih, and SuYu Yang. 2014. “Consumer Attitudes toward Online Video Advertising: An Empirical Study on YouTube as Platform.” IEEE.
Zaichkowsky, Judith Lynne. 1985. “Measuring the Involvement Construct.” *Journal of Consumer Research 12* (3): 341–352. doi:10.1086/208520.
Zhang, Kaifu, and Zsolt Katona. 2012. “Contextual Advertising.” *Marketing Science 31* (6): 980–994. doi:10.1287/mksc.1120.0740.
Zillmann, Dolf. 1971. “Excitation Transfer in Communication-Mediated Aggressive Behavior.” *Journal of Experimental Social Psychology 7* (4): 419–434. doi:10.1016/0022-1031(71)90075-8.