Check for updates

Consumer Neuroscience-Based Metrics Predict Recall, Liking and Viewing Rates in Online Advertising

Jaime Guixeres[1]*, Enrique Bigné[2], Jose M. Ausín Azofra[1], Mariano Alcañiz Raya[1], Adrián Colomer Granero[1], Félix Fuentes Hurtado[1] and Valery Naranjo Ornedo[1]

[1] *Instituto de Investigación e Innovación en Bioingeniería, Universidad Politécnica de València, València, Spain,*
[2] *Departamento de Comercialización e Investigación de Mercados, Facultad de Economía, Universitat de València, València, Spain*

The purpose of the present study is to investigate whether the effectiveness of a new ad on digital channels (YouTube) can be predicted by using neural networks and neuroscience-based metrics (brain response, heart rate variability and eye tracking). Neurophysiological records from 35 participants were exposed to 8 relevant TV Super Bowl commercials. Correlations between neurophysiological-based metrics, ad recall, ad liking, the ACE metrix score and the number of views on YouTube during a year were investigated. Our findings suggest a significant correlation between neuroscience metrics and self-reported of ad effectiveness and the direct number of views on the YouTube channel. In addition, and using an artificial neural network based on neuroscience metrics, the model classifies (82.9% of average accuracy) and estimate the number of online views (mean error of 0.199). The results highlight the validity of neuromarketing-based techniques for predicting the success of advertising responses. Practitioners can consider the proposed methodology at the design stages of advertising content, thus enhancing advertising effectiveness. The study pioneers the use of neurophysiological methods in predicting advertising success in a digital context. This is the first article that has examined whether these measures could actually be used for predicting views for advertising on YouTube.

Keywords: neuromarketing, YouTube, artificial neural networks, eye tracking, heart rate variability, brain response

INTRODUCTION

Advertising effectiveness is still challenging academics and practitioners. Neuroimaging and physiological measurement tools are becoming popular within marketing (Daugherty et al., 2016). Their primary uses are related to unconscious measures based on eye movement, heart rate and brain activity, among others (see Venkatraman et al., 2015 for more details). Such tools aim to provide better understanding of the impact of affect and cognition on memory (Vecchiato et al., 2013). Furthermore, neurophysiological methods can capture the dynamics of television commercials content because they provide continuous data unlike traditional measures, such as interviews and surveys that only reflect a global indicator for every commercial.

Global expenditure on media has been rising over the years and digital advertising is the fastest-growing category (McKinsey, 2015). Marketers are still calling for accurate assessments about advertisement effectiveness and about the return on advertising expenditure (McAlister et al., 2016). Nowadays, Internet advertising has evolved dramatically and a platform like YouTube is a good example on how to reach viewers at a global scale.

Advertisers pursue the attention of viewers and seek ad recall, brand recall and positive emotions. If this occurs, ads will be stored in the viewers' long-term memory. Humans may not remember each advertisement they have been exposed to, but neuroscience techniques can detect conditions that lead to the memorization of advertising.

The usage of neuro metrics in measuring advertising effectiveness overcome some of the weaknesses associated with traditional measures (Varan et al., 2015). Despite the benefits of using such measurements, the key question is the choosing of the variables of advertising effectiveness. Among the different types of effects pursue by advertising (see Moriarty et al., 2012), three types of effects have been considered. First, perception, and particularly exposure to the ad is recognized as the first step in any evaluation process. In this study, we adopt the online views as the measure of advertising effectiveness of online ads; secondly, the emotional dimension is typically used in evaluating the effects of advertising, thus we adopt the liking as an emotional metric; lastly, the cognition effect of advertising is measured through ad recall.

Neurophysiological methods offer richer data than self-report measurements of particular interest in advertising research. Firstly, physiological measurements of emotion allow researchers to analyze emotional activity without cognitive bias. Secondly, neurophysiological methods provide instant and continuous data that allow researchers to decompose the data analysis into small pieces of study. Lastly, physiological measurements typically offer a myriad of metrics. In the present study several metrics are compared and two new metrics derived from eye tracking (ET) are also proposed: "number of quadrants per second" (Quad_sec) and "gaze brand effectiveness ratio" (Brand_ratio). However, physiological measures have their own limitations: a strong reliance on physiological data to measure emotions leaves room for misinterpretation of physiological noise (e.g., natural changes in body status) and burdens researchers with the difficult task of attributing specific physiological changes (e.g., increase in heart rate) to complex and subjectively experienced emotions (e.g., hate, love, or fear).

Scholar research has recently adopted neurophysiological measures to better understand consumer responses to advertising (Astolfi et al., 2009). To the best of our knowledge, no articles in marketing have previously examined whether these measures could actually be transferred into real life views on advertising on YouTube. The work by Venkatraman et al. (2015) was one of the first pieces of research that tried to establish correlations between brand performance and physiological responses whilst viewers watched ads. So far it has been hard to gauge the number of viewers who will watch an ad, which in turn is one of the main objectives for marketers. However, digital platforms overcome this situation enabling researchers to measure consumer unconscious reactions to ads and the number of views.

Neurophysiological methods to measure advertising effectiveness are becoming popular including consolidated tools such as ET and facial reader (Wedel and Pieters, 2014), EGG (Ohme et al., 2010) and more recently sophisticated tools such as fMRI (Venkatraman et al., 2015; Couwenberg et al., 2017).

New extensions to sales are emerging, Thus, biomarkers can, to some extent, predict sales figures (see Kühn et al., 2016) and even in virtual reality experiments have been developed (see Bigné et al., 2016).

This paper aims to answer whether neurophysiological methods contribute anything beyond traditional methods in predicting ad success in a digital context. Specific research goals are listed below. Firstly, to analyze whether three of the most cited neurophysiological and behavioral techniques: electroencephalogram (EEG), heart rate variability (HRV) and ET correlate with common cognitive states typically used in advertising research (e.g., liking and recall measures) (Morin, 2011; Ruanguttamanun, 2014; dos Santos et al., 2015). Secondly, we aim to explain whether the variance in the number of views on a brand's official YouTube channel is related to any of the neurophysiological measures and their metrics.

A study to assess subjects' responses to nine 30-s online television ads was conducted. Data gathered from 47 subjects were split into six datasets based on three conditions: recall (RMB) vs. no-recall (FRG), liking (LIKE vs. DISLIKE) and Internet views (>5M vs. <5M). All the metrics extracted from physiological and behavioral responses were compared and correlated between these groups.

The contributions of this paper are listed hereafter. First, we show how different metrics from three neurophysiological devices correlate in an attempt to select the most accurate ones for digital commercials. Second, Artificial Neural Networks (ANN), using biometric data can predict digital views of ads, hence common physiological patterns related to unconscious responses can predict when an ad is going to be remembered or liked. Finally, two new metrics are proposed to measure advertising effectiveness in digital advertising, which show a high level of accuracy in predicting digital views. In this paper, we build on the later works that reviewed Super Bowl ads with ET and heart rate (Christoforou et al., 2015), and brain response (Deitz et al., 2016), but adding for the first time joined in neural networks models, three of the most employed signals in advertisement research (EEG, HRV and ET) and proposing two novel metrics based only on eye-tracking data that predicts viewer's preferences in video advertisements.

The rest of the paper is organized as follows. Firstly, we provide a brief literature review of advertising research effectiveness and we introduce neurophysiological methods related to ad recall and ad likeability. Secondly, we describe the experimental design and signal recording and processing techniques used to extract biometric data. Then, we describe the study results in three parts. In the first part, a comparison of the metrics from neurophysiological signals related to likeability of the ad and ad recall. In the second one, the correlations between these biometrics, the score given by participants in a poll (e.g., ACE_score) and the number of views on YouTube are exposed; and thirdly, by applying ANN to these biometric datasets, we predict the number of views on YouTube for each ad tested during the study. Finally, we discuss the contributions and implications for researchers and practitioners.

Established Methods in Advertising Research

Despite the diverse approaches used in advertising research (Vakratsas and Ambler, 1999), advertising success on ad execution has focused on traditional measures such as liking, excitability, and recall (Venkatraman et al., 2015). Acknowledging the established literature (Astolfi et al., 2009; Kim et al., 2014), this paper focuses on liking and recall as traditional measures. Online polls have been adopted in academic research as a valuable data source (see Strach et al., 2015). Based on a US national-representative Internet sample of 500 respondents, Ace Metrix has been providing advertising effectiveness scores since January 2009 and it is used in this study.

Advertising Research in a Digital Setting

Digital channels have changed advertising research dramatically and, as a result, a new paradigm is emerging (Ha, 2008; Bigné, 2016). One of the major gains is that analytics are available at ad level, including exposure measured through number of views and likeability through "likes."

Neurophysiological Tools in Advertising Research

This study focuses on three neurophysiological methods, aiming to collect data from different angles: eye movements, heart variability and brain responses.

Eye tracking is a well-established measure of visual attention (Wedel and Pieters, 2014; Venkatraman et al., 2015) to different stimuli, such as product choice (Guerreiro et al., 2015), static images (Mould et al., 2012), printed ads (Elsen et al., 2016) banner ads (Lee and Ahn, 2012) and videos of the Super Bowl (Christoforou et al., 2015).

Heart rate variability is the physiological phenomenon of variation in the time interval between heartbeats. It is measured by the variation in the beat-to-beat interval (Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology, 1996). This variability of the heart is related to activations of the sympathetic and parasympathetic systems of the autonomic nervous system. HRV provides an independent measure of attention (Lang et al., 1999) and it has been applied to television commercials (Acharya et al., 2006; Grandjean et al., 2008; Geisler et al., 2010; Bellman et al., 2013; Valenza et al., 2014; Venkatraman et al., 2015).

Electroencephalogram is an electrophysiological monitoring method to record the electrical activity of the brain. The relationship between affection, engagement and brain activation in frontal brain activity has been well documented in psychology and neuroscience research (Harmon-Jones et al., 2010; Khushaba et al., 2013). Emotional frontal asymmetry as hypothesized by Davidson (2004) has been applied to analyze commercials (Ohme et al., 2010; Vecchiato et al., 2011), including Super Bowl ads (Deitz et al., 2016) and advertising success (Venkatraman et al., 2015). Vecchiato et al. (2011, p. 582) showed that "activity" in the left-frontal cortex related to "pleasant" commercials and activity in the right-frontal cortex associated with "unpleasant" commercials.

Hypothesis Development

As briefly discussed earlier, ET, HRV and EEG provide measures of responses to advertising stimuli and might be related to ad performance. This study attempts to examine the relationship between three types of data from neurophysiological tools, ET, HRV and EEG, and three advertising variables typically used in advertising research, such as ad recall, ad likeability and ad views. Hypotheses will be anchored in three streams of research aiming to integrate them into a single approach: (i) theoretical advertising literature; (ii) online advertising; (iii) neurophysiological research related to advertising. An integrative approach is useful because neurophysiological primary data *per se* are non-meaningful for advertising research. Therefore, this type of data must be interpreted in relation to classic advertising assumptions in order to prove their validity. Most of the data gathered in this type of studies are based on a different methodological paradigm that derives from psychophysiology (Bolls et al., 2012).

1. Advertising research shows a positive relationship between recognition and attention. Donthu et al. (1993) found a positive relationship between attention and recall for outdoor advertising. In an ET study, Pieters et al. (2002) found that attention led to ad recognition. Baack et al. (2008) posit that recognition is considered an immediate measure of attention.
2. As discussed earlier, ET and some metrics from EEG can capture attention, hence if attention leads to recall, it could be argued that each neurophysiological tool can capture attention.

The Advertising Research Foundation's Copy Research Validity Project (CRVP) showed in the early nineties that advertising likability is the single best measure of effectiveness (Rossiter and Eagleson, 1994). Furthermore, likeability has been considered relevant and important in measuring commercial effectiveness in the ads aired in the Super Bowl, showing stable scores between 1990 and 1999 (Tomkovick et al., 2001). The positive influence of likeability has also been highlighted recently in online settings. Thus, likeability of online video ads has successfully linked to intention to share them (Shehu et al., 2016), which can be interpreted as a successful performance. Recent literature in neuromarketing also highlights a relationship between liking and HRV, ET measurements and fMRI signals (Venkatraman et al., 2015).

Online video platforms, such as YouTube, have been largely approached from the user-generated content perspective (see Smith et al., 2012). However, its dimension as a digital channel for watching commercials has almost been neglected, with some related exceptions (Verhellen et al., 2013). The number of views of each online video, including commercials, is available on YouTube, and it is commonly seen as a valid measure of its popularity. Given its social media nature, recent research is addressing two main fields of interest: the sources that drive views to a video and the preferred type of content. A recent study by Zhou et al. (2016) identified YouTube search and related video recommendation as the major view sources. In

adopting YouTube views, the age of the video and the potential replays must be considered. Research shows that user's preference seems relatively insensitive to the video's age (Cha et al., 2007). More recently, Chen et al. (2014) analyzed a lifetime model of online video popularity that features the following three main characteristics of interest here, adopting views as a potential variable in explaining our intended relationships: (i) views follow a Zipf distribution; (ii) replay percentage is very low; (iii) and only video content on news and sports are strongly dependent on age, with popularity being much less sensitive to age in music videos, which can be considered closer to ads. Therefore, views can be adopted as a valid measure over time of ad exposure.

Based on previous reasoning on recall, likeability and number of views, and their relationship with neurophysiological tools, therefore:

– H1: Eye tracking, HRV and brain activity capture (a) recalled and non-recalled online ads, (b) liked and disliked online ads, and (c) ads with high numbers vs. low numbers of YouTube views.

1. Online polls (e.g., Ace Metrix, YouGov) are becoming popular in the industry and have been used for academic purposes. Established scores in the market are considered by marketers as a valuable source of information. Ace Metrix is a company that tests ads within 24–48 h of each ad's initial airing and provides meaningful scores. Ace Metrix component scores range between 1 and 950. Scores for the general population are normally distributed and approximately centered around 530. For obtaining Ace Score, every ad is shown to a unique set of 500 respondents who complete standardized surveys that assess the ad both quantitatively and qualitatively providing rich insights by demographic segment. All ads are scored on 6 factors most likely to influence consumer behavior including relevance, likeability, information, change, attention and desire—as well as re-watch ability, purchase intent and brand linkage. After respondent scores are collected, Ace Metrix component scores are computed and assigned to each ad, creating an overall score called the Ace Score. Since its inception, Ace Metrix has consistently used the same methodology to measure the effectiveness of every ad they have tested. As a result, they are in a unique position in the field of advertising effectiveness to assess relative advertising performance between any competitive set of ads imaginable, both across different industries and different time periods.

2. In addition, the digital world is changing data availability and triggering new analytical research methods where new research avenues for unstructured data, including neurodata, and new methods like ANN call for new research (Wedel and Kannan, 2016). Venkatraman et al. (2015) conducted a similar study but in the offline context. They found that deceleration correlated with liking ($r = 0.37$, $p < -05$) and recognition ($r = 0.34$, $p < 0.05$). The current study differs in two ways: (i) we parse neurophysiological correlations in a different context,

such as digital exposure featured by searching rather and displaying; (ii) Venkatraman et al. (2015) captured advertising effort, through GRPs, and advertising outcome, through advertising elasticities; however, our study attempts to find out the correlates of neurophysiological metrics from ET, HRV and EEG with independent variables of ad effectiveness based on both, survey data and digital views. Therefore, we predicted:

– H2: Eye tracking, HRV and brain activity capture correlate with (a) self-reported score of ad effectiveness and (b) online views on YouTube.

As stated before, YouTube as a digital channel to watch commercials has almost been neglected with some related exceptions (Verhellen et al., 2013). Our aim here is to use ANN to predict the number of online views of ads placed in YouTube, where the input variables are metrics from ET, HRV, and EEG. ANN is useful for parsing non-linear relationships and adopt feed forward and back propagation approaches (West et al., 1997). ANN has been successfully applied in advertising since the mid-nineties (Curry and Moutinho, 1993). Research posits the superiority of these methods over other statistical approaches. Surprisingly, this is the first attempt to use ANN to predict online views. In our study, we aim to classify and to predict online views based on metrics from ET, HRV, and EEG. Therefore, we predicted:
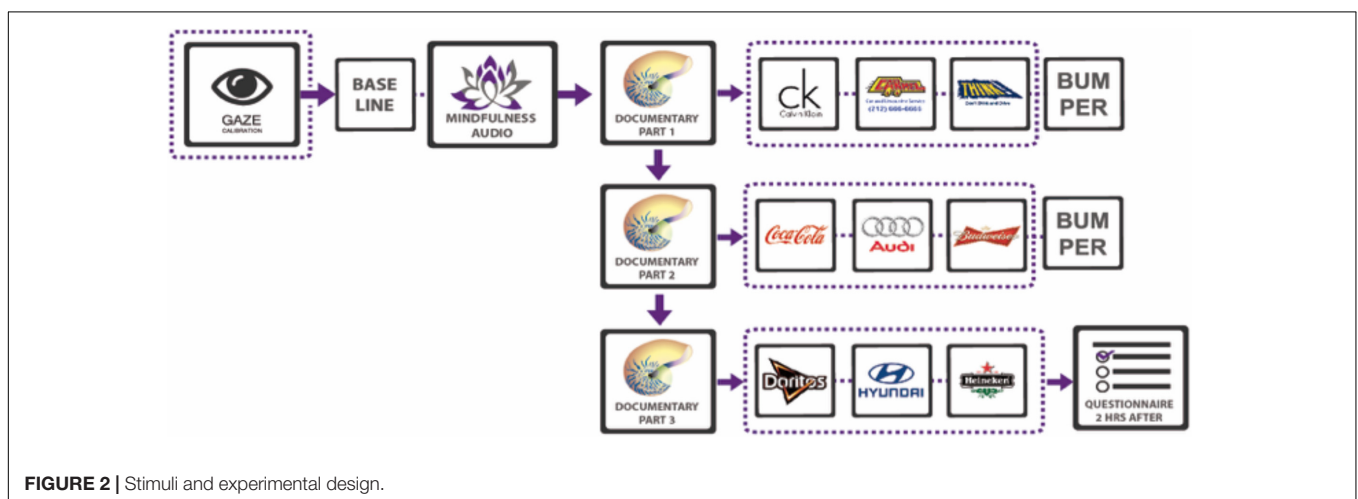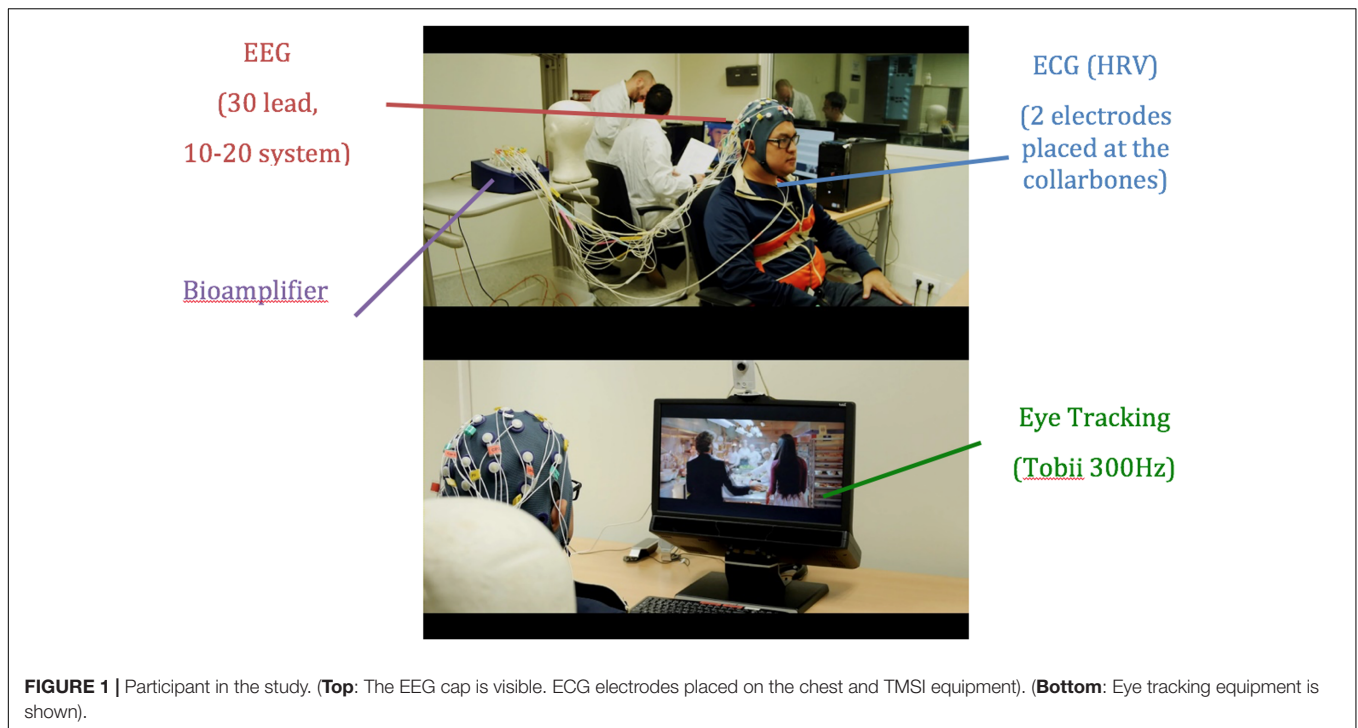
– H3: Artificial neural networks based on mixed data from ET, HRV and brain activity predict the number of online views on YouTube.

MATERIALS AND METHODS

Participants and Design

Final sample consisted of 35 randomly healthy volunteers (15 women and 20 men, mean age = 25; $SD = 5$ years) recruited from the city where the lab is located. Initial sample measured was 47 subjects but after an examination of the dataset was carried out, 12 participants were removed due to corrupted data from experimental sessions in some of the acquired signals. All of the participants showed corrected-to-normal vision and hearing. They were asked to pay attention to the documentary as in a common situation. No mention of the importance of the ads was made. The study was approved by the Institutional Review Board of the *Polytechnic University of Valencia* with written informed consent from all subjects in accordance with the Declaration of Helsinki.

The experiment was conducted in a neuromarketing lab of a large European university and comprises the three parts shown in **Figure 1**. In Parts 1 and 2, participants sat comfortably on a reclining chair with a 32-channel EEG device, with two electrodes to measure heart variability and an eye-tracker (**Figure 1**). In Part one, participants were exposed to a mindfulness audio designed by experts to help them relax and disconnect from past experiences of the day (Fjorback et al., 2011; Demarzo et al., 2014). Then, in Part two they were shown a 30-min long

FIGURE 1 | Participant in the study. (**Top**: The EEG cap is visible. ECG electrodes placed on the chest and TMSI equipment). (**Bottom**: Eye tracking equipment is shown).

FIGURE 2 | Stimuli and experimental design.

documentary with three commercial breaks of three ads lasting about 30 s each; the first break occurred after 7 min, the second in the middle of the documentary, and the third 7 min before the end as **Figure 2** depicts. At the end of this second part, participants were informed that an interview would be held 2 h later (Part 3).

Pretesting and Stimuli

Television commercials from the final of the Super Bowl 2015 were chosen because they are a good representation of the most searched high impact ads. We wait for a year to get results of number of views on the brand's official YouTube channel and a final selection of eight from 47 ads was made to represent a uniform distribution of ads ranked by number of views and

also to represent a distribution of different commercial products (**Table 1**). In addition to the number of views, the ACE Metrix[1] score was also obtained for some ads. The ACE Metrix is the most employed US scale drawn up by consumers that evaluates the ad creative effectiveness based on viewer's reaction to national TV ads. The results are presented on a scale of 1–950. The selected television commercials belong to international brands of commercial products such as drinks (3), food (1), cars (2), textiles (1) and services (1). None of the ads had been broadcasted in the country where the experiment was performed in order to remove previously unchecked exposure of the subjects to the proposed stimuli. These videos were randomly distributed

---

[1]http://www.acemetrix.com/

during the sessions with participants to avoid bias in data analysis.

Data Recording and Processing
Cerebral Recording (EEG)
Electrical activity of the brain was recorded by a stationary 32-channel system (REFA 32, TMSI hardware). EEG activity was gathered at a sampling rate of 256 Hz. The experiment used 30 Ag/AgCl water-based electrodes and bracelets attached to the opposite wrist of the subjects' dominant hand. The montage of brain electrodes followed the international 10–20 system (Jasper, 1958).

The EEG baseline was removed and channels detected as having corrupted data were rejected and interpolated from the closest electrodes (Colomer Granero et al., 2016). When a channel with erroneous data was identified, kurtosis was employed computing the fourth standardized moment in the signal of each electrode. This kurtosis is defined in Equation 1

$$K(x) = \frac{\mu_4}{\sigma_4} = \frac{E[(x - \mu)^4]}{E[(x - \mu)^2]^2} \qquad (1)$$

where μ_4 was the fourth moment of the mean, σ was the standard deviation and $E[x]$ was the expected value of signal x. The EEG signal was segmented in one second acquired accordingly to the experiment events. The intra-channel kurtosis level of each epoch was used to reject the epochs with high levels of noise.

To detect artifacts from eye movements, blinking and muscular activation, Independent Component Analysis (ICA) (Gao et al., 2010) and automatic method (ADJUST) (Mognon et al., 2011) were implemented. Each EEG artifact-free trace was band pass filtered twice in order to isolate only the spectral components in delta (1–3 Hz), theta (4–7 Hz), alpha (8–12 Hz), beta (13–24 Hz), beta extended (13–40 Hz) and gamma bands (25–60 Hz).

To quantify the cerebral activity in each band, the Global Field Power (GFP) (Wackermann et al., 1993) was calculated as explained in previous work (Colomer Granero et al., 2016).

Recent studies have shown that the main areas involved in the phenomena of memorization and pleasantness are the frontal areas (Astolfi et al., 2009). For that reason, calculation

TABLE 1 | ACE metrix score, number of visits on a brand's official YouTube channel during a year and ranking of visits established for the selected video advertisements regarding number of visits.

	ACE_score	No. visits	Rank of visits*
Drink 2	665	12,037,414	D
Car 2	394	10,583,315	D
Drink 3	Non-evaluated	10,356,789	D
Drink 1	641	6,652,069	C
Food	626	4,594,507	B
Textile	362	2,952,014	B
Car 1	611	1,453,022	B
Service	167	15,340	A

*VISIT_RANK: A: = <1M views; B = 1M–5M views; C = 5M–10M views; D = >10M views.

of electrodes in frontal lobe were taken into account. A GFP signal was then calculated for each frequency band considered in the experiment. GFP associated with each ad analyzed and GFP during a period taken as the baseline of watching a 2-min neutral documentary before the block of ads were compared and normalized to obtain the corresponding z-score index.

In addition to the z-score of GFP for each EEG band, two metrics applied in advertising research were also calculated: The *Pleasantness Index (PI),* and the *Interest Index (II).* The PI is a metric calculated over time that provides information about the pleasantness of the stimuli presented (Vecchiato et al., 2011). Brain activity gathered by left-frontal electrodes is compared with brain activity registered by the right-frontal electrodes (frontal asymmetry). These comparisons are made with GFP in the theta and alpha bands, comparing asymmetric pairs of electrodes.

The questionnaire generated the likeability score for each ad under study. Using this information, participants were segmented into two groups: "LIKE" and "DISLIKE." Then, the brain Pleasantness Index, PI was calculated for each group as describes (Colomer Granero et al., 2016).

The II, enables an advertising evaluation of user interest in theta and beta bands (Vecchiato et al., 2010). The most relevant peaks of these signals are selected. Two parameters were obtained: the number of peaks during a particular ad (PN_{total}) and the number of peaks during the periods the brand name appeared in that particular ad (PN_{brand}). Accordingly, the II was calculated as describes (Colomer Granero et al., 2016).

Heart Rate Variability
To analyze HRV, the electrocardiogram, ECG signal needs to be filtered (Blanco-Velasco et al., 2008), analyzed to detect QRS zones (Pan and Tompkins, 1985) and revised manually by an expert, because the appearance of a single ectopic can produce variations in certain key parameters extracted from this analysis (Clifford, 2006). HRV analysis can generate a set of metrics that can be extracted from different dimensions: time, frequency, time-frequency and non-linear. Parameters extracted from the time domain used in this study were: average heart rate (t_meanHR), standard deviation of continuous HR values (t_sdHR), the square root of the sum of successive differences between adjacent RR intervals (t_RMSSD) and the number of successive pairs of RR intervals showing a difference of more than 50 ms between them (t_NNx).

Power spectral density (PSD) analysis provides information about the amount of power in the frequency bands defined for the beat-to-beat interval signal generated. In this case, we employed the Lomb-Scargley method (Castiglioni and Di Rienzo, 1996). The frequency bands defined (low frequency, LF and high frequency, HF) are stated by Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology (1996). Power metrics can be presented in absolute values (aLF, aHF, aTotal), normalized to total energy (nLF, nHF) or in a percentage value of total energy (pLF, pHF). The ratio established between the LF and HF band provided information on the sympathetic/parasympathetic balance. The power value of the peak in the fundamental frequency (peakLF, peakHF) was also extracted.

Non-linear analysis was run using techniques such as the Poincaré graph that give us SD1 and SD2 metrics (Fishman et al., 2012). Sample entropy (sampen) is another non-linear technique that attempts to quantify the complexity or degree of new information generated (Richman and Moorman, 2000). If entropy is equal to zero, then consecutive sequences are identical. Similarly, bigger values show higher complexity of the analyzed signal.

To summarize, for each ad analyzed by subject, HRV metrics were computed by means of a computational analysis plug-in based on Matlab (Guixeres et al., 2014). For the purpose of this study, the seventeen most relevant metrics employed in HRV analysis (Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology, 1996) were selected based on time, frequency and non-linear domains.

Eye Tracking

The Tobii TX300 eye tracker[2] was used in this experiment as **Figure 1** depicts. This eye-tracker collects gaze data at 300 Hz. The subsequent analysis of raw data used Tobii Studio 3.2 software. For each commercial the following metrics were obtained from the gaze data: (i) number of fixations during each ad (Fix_Count_Advert); (ii) average duration of fixations during an ad (Fix_Dur_Advert); Furthermore, we obtained several metrics from the times the brand appeared in each ad. To calculate such metrics, a dynamic Area of Interest (AOI) that followed the brand was created using TOBII studio software to obtain (iii) the average duration of fixations exclusively focused on the brand (Fix_Dur_Br); (iv) the number of fixations during the brand appearance that focused on it (Fix_Count_Br); (v) the time from the appearance of the brand until it was fixated on for the first time (FFIX_Dur_Br); (vi) the number of visits inside the brand's AOI (Visit_Count_Br); and (vii) average duration of visits to the brand's AOI (Visit_Dur_Br). For the last two metrics, it should be remembered that a visit is the event that starts when the eye enters an AOI until it leaves such AOI. In addition to these metrics, two new metrics are proposed in this study: (viii) Number of Quadrants per second (Quad_sec); and (ix) the Gaze Brand Effectiveness Ratio (Brand_ratio).

Quad_sec enables the way the user explores a space with his eyes to be quantified. To calculate this metric, the screen surface was divided into a grid of 4 × 4 equal-sized quadrants. Then, the average quantity of different quadrants that the eye visited per second was calculated for each commercial. Higher values for this metric meant that the subject explored the space in "ambient mode," covering all the space with his eyes, whilst lower values meant that the subject explored the space in "focus mode," centering his visual attention on specific zones. These two modes of watching an image stimulus have been reported in previous works (Bradley et al., 2011; Holmqvist et al., 2011).

$$\text{Quad}_{\text{seg}} = \frac{N_q}{t_s} \qquad (2)$$

---

[2]Tobii.com

where N_q is the number of visits to quadrants during stimulus presentation and t_s is the duration in seconds of the stimulus.

Brand_ratio enables the effectiveness of visual attention toward the brand during the ad to be quantified. To calculate Brand_ratio, brand appearance was controlled by setting an AOI around the brand every time it appeared in the commercial. Then this metric was defined as the number of seconds that the subject looked directly at the brand divided by the total time that the brand was present on the screen during the commercial. This metric could be related to the participant's interest in and familiarity with the brand as it relates to the time that eye and brain are able to identify a brand, a concept related to familiarity (Kent and Allen, 1994).

$$\text{Brand}_{\text{ratio}} = \frac{t_{bf}}{t_b} \qquad (3)$$

where t_{bf} is the time in seconds that the gaze fixed on the brand and t_b is the total time in seconds that the brand appeared during the ad.

Questionnaire

Data were sorted using three criteria: (i) spontaneous ad recall; (ii) ad liking; (iii) the number of online views was used as a control variable. The criterion for ad recall was to remember without clues brand names of the commercials 2 h after the study. Accordingly, participants were split into two subgroups. The first dataset was related to the biometric activity collected during the observation of the recalled commercials 2 h after being exposed to them. This dataset was named RMB. The second subset included the biometric activity collected during observation of the non-recalled commercials, (FRG). The ad liking criterion was related to the biometric activity collected during observation of the television commercials that the subjects rated 5 or above on a 10-point Likert scale, being this subset named LIKE and DISLIKE, respectively.

Statistical Methods

In order to test comparisons of means of the metrics calculated, a set of Shapiro-Wilk tests (W) were conducted to test whether dependent variables deviated from normality. Then a statistical analysis was carried out using ANOVA for metrics with normal distribution and the Mann Whitney non-parametric test for metrics that did not show a normal distribution. A corrected p-value less than $p = 0.005$ was chosen to correct multiple comparisons effect (Feise, 2002).

To get the correlation of neurometrics with the ACE score and the number of online views of the ads on YouTube, a Pearson correlation was applied to the number of visits so it could be considered as a linear scale. However, Spearman's correlation was applied to the ACE score so it could be considered as a rank variable instead of a linear scale.

Artificial Neural Networks

We adjusted two neural networks with SPSS statistics using all metrics defined for EEG, HRV and ET, and including gender, as our input variables. The first network was adjusted to classify advertising responses into a ranking for the number of views

on YouTube (RANK_VISITS). This ranking divided ads into four clusters (<1M: ads with less than 1 million visits, 1M–5M, 5M–10M, and >10M: ads with more than 10 million visits). The second network was adjusted to predict the real number of visits for each ad.

Two kinds of network were compared with the same data. Multi-layer Perceptron networks (MLP) and Radial Basis Function networks (RBF). Regarding accuracy for classification and estimation of the data, MLP networks were selected finally for the two purposes instead of RBF. After testing results changing several parameters in MLP architecture, a final structure was chosen for the both neural networks (see Appendix). To validate accuracy of networks, cross-validation technique was employed. Entire sample was divided into two groups (70% of cases for training the network and 30% of cases to assess classification accuracy) and that validation was repeated 10 times ($k = 10$), selecting each time different groups for training and assessing. Final results were averaged from the 10 turns.

RESULTS

Biometric Mean Comparison
In order to test H1, we conducted a comparison among means of the metrics calculated from EEG, HRV and ET. As stated earlier, these metrics were compared by means of the following two conditions, recall vs. non-recall after 2 h (RMB vs. FRG) and by likeability (LIKE vs. DISLIKE).

Brain Response Comparison
Figure 3 shows the comparison among means for z-score indexes in each frequency band for the different factors chosen. In the case of remembered ads, the RMB group shows significant differences compared to the FRG group with higher values in the delta, theta, beta ext. and gamma bands. The LIKE group showed significant differences compared to the DISLIKE group with higher values in the delta, theta, beta ext. and gamma bands.

In the pleasant index (PI) and II, there were no significant differences in the comparison between the recalled and liked groups.

HRV Response Comparison
Table 2 shows the comparison of means for HRV metrics for the different factors chosen. In the case of the remembered ads, the RMB group showed significant differences compared to the FRG group, with higher values in the non-linear SD2 Poincaré index (p_SD2) that reflect higher continuous beat-to-beat variability (Piskorski and Guzik, 2007). The LIKE group showed significant differences compared to the DISLIKE group with higher values in the energy of the low frequency band (f_aLF_lomb) which is associated with sympathetic activation (Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology, 1996).

ET Response Comparison
Table 3 shows the comparison of means for eye-tracking metrics for the different classic metrics and the two-new metrics

proposed. In the case of the recalled ads, the RMB group showed significant differences vs. the FRG group, generating lower values in (Visit_Dur_Br). The LIKE group did not show significant differences compared to the DISLIKE group.

Hypothesis 1 is confirmed, as there are significant differences in each signal (EEG, HRV and ET) between ads (i) recalled and non-recalled (22% of comparisons) and (ii) liked and disliked (16% of comparisons).
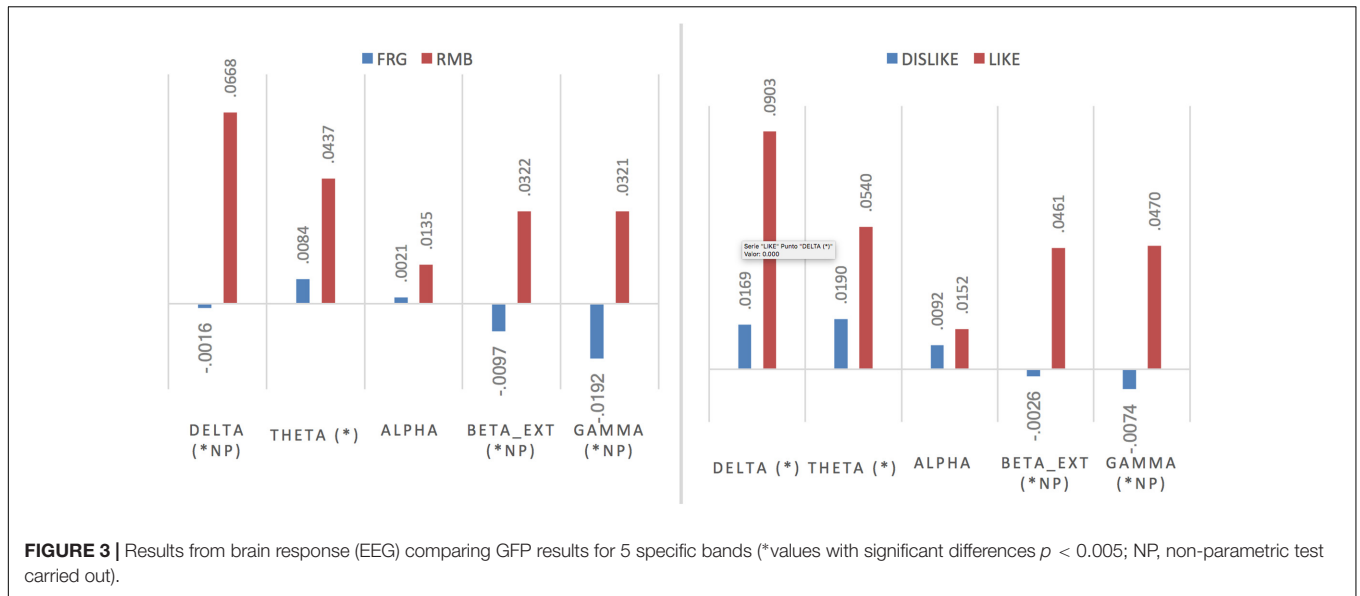
As regards H2, centered on the correlation of neurometrics with the ACE score and the number of online views of the ads on YouTube, 19 of the 25 metrics showed significant ($p < 0.01$) correlation with the ACE score and 15 of the 25-metrics showed significant correlation with the number of online views. In EEG, the z-score in the delta band correlated with both indexes and the z-score in the theta band correlated with the ACE score. Both indexes, pleasantness and interest, showed high values in terms of significant correlation with the ACE score. In particular, there was a high level of correlation between PI_theta and the number of visits. In terms of ET, the proposed metric (Brand_ratio) showed significant correlations with both indexes, PI and II. In addition, Quad_sec showed significant correlation with the ACE_score. Visit duration and Fixation count on the brand showed significant correlation with the number of visits. Fixation duration showed a negative correlation with the ACE score and the value of the correlations of the Fixation count during the ad for both indexes was relevant. Focusing on HRV, t_NN50 showed significant correlations with both indexes. The total energy band; LF and HF bands, also showed correlations with both indexes. The normalized values of LF and HF and the sympathovagal index (LFHF) showed significant correlations with the ACE score. Frequency with the maximum peak on the HF band (peakHF) and the SD2 Poincaré index also showed significant correlations with the number of visits. The type 1 non-linear parameter sample entropy index (sampen1) showed significant correlations with both indexes whilst sampen2 only correlated with number of visits. Therefore, H2 is confirmed, showing significant correlations between metrics for EEG, HRV and ET with ACE score and number of visits on YouTube.

Predicting Ad Effectiveness on the Internet with ANN
ANN to Rank Visit Classification
Two kinds of ANNs were compared with the same data: MLP and RBF. In terms of accuracy when classifying data, MLP networks worked better than RBF. After testing results and changing several parameters in the MLP architecture, the final structure chosen had one hidden layer with a tangent hyperbolic activation function and an output layer with a softmax activation function. All the co-variables were typified (see Appendix for more details).

Table 4 shows the final results for the classification of the training and the test dataset using the neural network. The percentage of correct predictions in the test dataset was 82.9%. Three higher ranking levels (1M–5M, 5M–10M, and >10M) showed high effectiveness ratios of classification (88.5, 100, and 90%, respectively). Only the first group (<1M) showed a poor

FIGURE 3 | Results from brain response (EEG) comparing GFP results for 5 specific bands (*values with significant differences $p < 0.005$; NP, non-parametric test carried out).

effectiveness ratio of 36.4%, mistaking more than 54% of cases for second level (1M–5M).

Of the 37 input variables used in neural networks, the importance of each normalized metric to predict was extracted from the neural network. This metric shows the most relevant variables for classifying each case according to the correct ranking. The pleasantness index extracted from the theta band (PI_theta, 100%) was the most important parameter for brain metrics, followed by II_theta (56.6%) and PI_ alpha (56.5%). Regarding HRV metrics, the mean of HR (64.20%) was the most representative parameter, followed by SD1 Poincaré (61.40%) and

t_RMSSD (58.30%). In terms of eye-tracking metrics, the fixation count during the ads (59.30%) was the most important index, followed by the average fixation count during brand appearance (59.30%), and the average duration of fixations during brand appearance (45.30%). Gender inclusion and whether the ad was consciously remembered came last in the ranking with very little importance. The proposed metrics from ET showed medium importance (Ratio_Brand: 40.80% and Quad_sec: 40.20%). When comparing the frequency brain bands, the theta band behaved best, in the PI index (100%), II index (56.6%) and z-score (37.50%).

TABLE 2 | Mean comparisons of HRV metrics in the time, frequency and non-linear domains.

HRV Metrics	RMB		FRG			UNLIKE		LIKE		
	Mean	SD	Mean	SD	p	Mean	SD	Mean	SD	p
t_meanHR	71.9	12.3	70.6	8		69.6	8.8	71.1	7.8	
t_sdHR	5.2	7.3	5.5	4.5		5.1	5.6	5.3	2.7	
t_NN50	13.3	18.4	13.8	12.7		13	13.6	13.5	13.3	
t_RMSSD	43.5	32.3	54.6	112.3		43.2	26	60.1	142.7	
f_aLF_lomb	0.031	0.015	0.033	0.013		0.031	0.014	0.035	0.014	(*NP)
f_aHF_lomb	0.027	0.020	0.027	0.017		0.029	0.021	0.026	0.017	
f_aTotal_lomb	0.060	0.027	0.061	0.022		0.062	0.028	0.062	0.024	
f_nLF_lomb	0.539	0.223	0.573	0.2		0.547	0.203	0.594	0.194	
f_nHF_lomb	0.461	0.223	0.427	0.2		0.453	0.203	0.406	0.194	
f_LFHF_lomb	2.022	2.156	2.121	2.057		1.786	1.545	2.412	2.412	
f_peakLF_lomb	0.1	0.02	0.09	0.02		0.09	0.02	0.09	0.02	
f_peakHF_lomb	0.27	0.07	0.26	0.07		0.26	0.07	0.25	0.07	
p_SD1	30.7	22.8	32.8	21.4		30	18	32.9	19.8	
p_SD2	66.1	28	77.6	36.7	(*NP)	68.8	30.4	80.3	38	
p_SD1vsSD2	0.497	0.382	0.441	0.192		0.462	0.224	0.423	0.154	
nl_sampen1	2.175	0.227	2.141	0.243		2.179	0.25	2.143	0.217	
nl_sampen2	1.803	0.486	4.19	0.532	(*NP)	9.373	0.788	1.66	0.389	

*Values with significant differences $p < 0.005$; NP, non-parametric test.

ANN for Estimating the Number of Online Views

As described above, MLP and RBF were compared only in terms of accuracy (time execution was extremely low in both cases). MLP networks also worked better than RBF networks in terms of accuracy when estimating the number of visits. After testing results and changing several parameters in the MLP architecture, the final structure chosen had one hidden layer with a tangent hyperbolic activation function and an output layer with an identity activation function. All the co-variables and the dependent variable were typified.

For the final MLP structure, the first subset was used to train the MLP and the estimated values of the number of visits were obtained from the second subset to test the accuracy of the network. Final results for the estimation of training and the test dataset using an ANN. Relative error from the test dataset was 0.199, that is, a significant level of variance.

Of the 37 input variables used in the neural networks, the pleasantness index extracted from the theta band (PI_Theta, 100%) was again the most relevant variable in the neural network in predicting the number of views. Regarding brain

metrics, PI_theta was followed by PI_alpha (13.10%) and II_theta (11.60%), though these were a long way behind. In HRV metrics, sample entropy (53.70%) was the most important parameter, followed by the total energy of frequency band (43.40%) and t_RMSSD (41.70%). Regarding eye-tracking metrics, the count of fixations during ads (44.30%) was the most important index, as on the first occasion, followed by number of visits during brand appearance (27%) and then followed very closely by the count of fixations during brand appearance (26.60%). The inclusion of gender in the model and whether the ad was consciously recalled ranked last and do not improve prediction. The proposed new metrics extracted from ET showed different results. Quad_sec showed a medium-low importance (16%) and Ratio_Brand a low importance (6.70%). Half the HRV metrics were in the upper positions.

Figure 4 represents predicted vs. observed values as a scatterplot for each case. Figure shows that neural network classifier worked well in classifying real data regarding biometric response, specifically observing the cluster of ads with higher number of views vs. ads with poor audience. There was some dispersion in all the ads in terms of real values.

**TABLE 3 |** Mean comparisons of ET metrics.

	RMB		FRG			UNLIKE		LIKE		
Eye Tracking Metrics	Mean	SD	Mean	SD	p	Mean	SD	Mean	SD	p
Brand_ratio	0.36	0.25	0.32	0.25		0.37	0.26	0.31	0.24	
Visit_Count_Br	6.7	2.42	6.65	2.33		7.29	2.53	6.32	2.38	
Visit_Dur_Br	0.76	0.24	0.63	0.25	(*NP)	0.75	0.23	0.62	0.38	
FFix_Dur_Br	0.18	0.14	0.16	0.13		0.17	0.13	0.15	0.13	
Fix_Dur_Br	0.18	0.12	0.17	0.11		0.18	0.15	0.17	0.12	
Fix_Count_Br	28.23	28.97	23.73	22.87		33.76	37.48	21.98	19.29	
Quad_sec	2.47	1.35	2.48	1.29		2.65	1.62	2.58	1.24	
Fix_Dur_Advert	0.24	0.1	0.24	0.11		0.24	0.12	0.21	0.11	
Fix_Count_Advert	273.32	120.34	281.8	200.12		340.73	180.73	270.83	165.43	

Values with significant differences p < 0.005; NP, non-parametric test.

TABLE 4 | Results for the successful classification of training and test dataset applying an adjusted neural network.

		Classification				
	Real rank order	Predicted				
		<1M	1M–5M	5M–10M	>10M	% succ.
Training	<1M	17	1	0	0	94.4%
	1M–5M	0	60	0	0	100.0%
	5M–10M	0	0	23	1	95.8%
	>10M	0	0	0	67	100.0%
	Overall percentage	10.1%	36.1%	13.6%	40.2%	98.8%
Test	<1M	4	6	1	0	36.4%
	1M–5M	2	23	0	1	88.5%
	5M–10M	0	0	9	0	100.0%
	>10M	0	0	3	27	90.0%
	Overall percentage	7.9%	38.2%	17.1%	36.8%	82.9%

Number of cases in each possible class and percentage of successful cases.

FIGURE 4 | Scatterplot showing predicted vs. observed-values for number of visits.
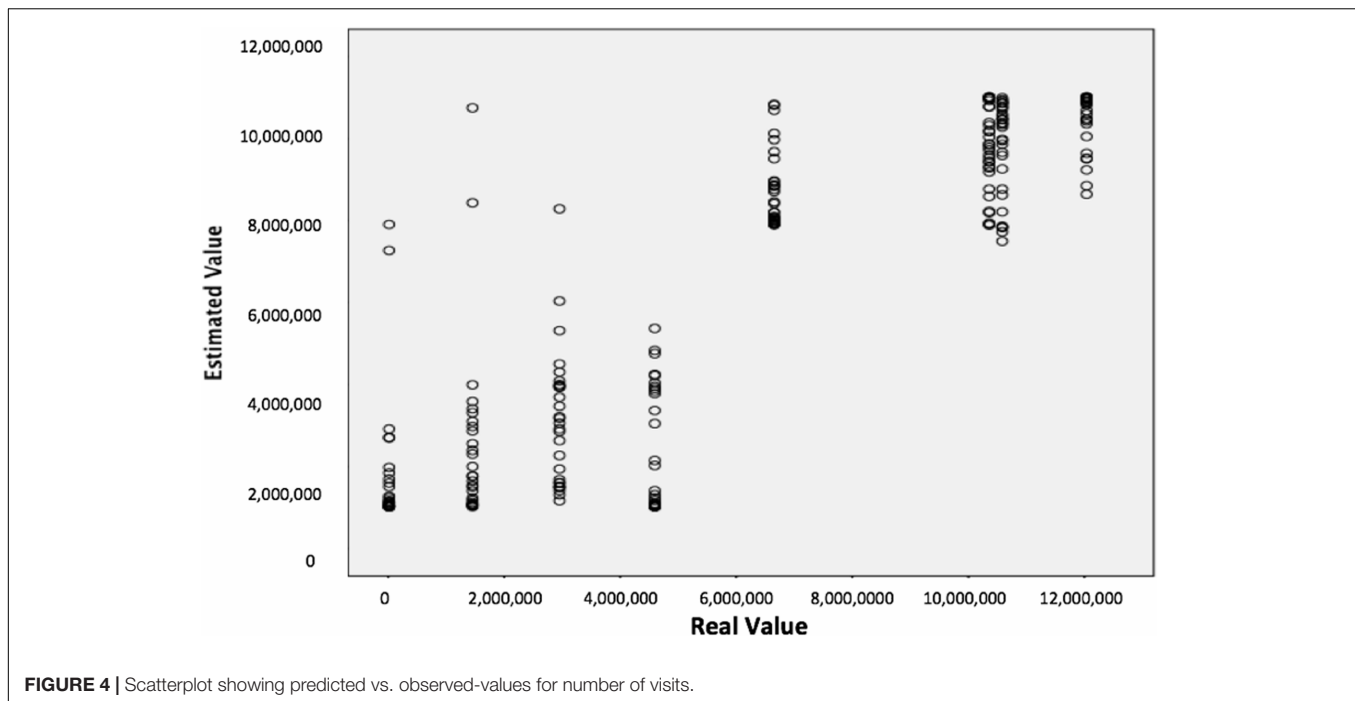
DISCUSSION

In this study, several hypotheses have been confirmed. The first hypothesis was supported, showing significant differences in physiological responses (EEG and HRV) and ET for each of the three dimensions raised (i.e., recall, liking and visits). In cases where participants remembered the ad after 2 h, results showed higher probability that the spectral amplitude in the RMB condition was always higher than the power spectra in the FRG conditions (Astolfi et al., 2008). A statistical increase of PSD in the prefrontal and parietal areas for the RMB dataset compared with the FRG dataset was in line with the suggested role of these regions during the transfer of sensory perceptions from short-term memory to long-term memory storage. Specifically, there were higher values in the theta band for the cases where the ad was remembered, which is in line with other studies (Werkle-Bergner et al., 2006; Boksem and Smidts, 2014; Vecchiato et al., 2014). Regarding this condition, an interesting future analysis should look at the biometric response differences between people that remember the ad without remembering the brand or vice versa.

In terms of the HRV analysis, the means of the sample entropy parameters were significant, showing the higher complexity of heart variability in cases where the ad was remembered. Valenza et al. (2012), showed that Approximate Entropy, decreased during arousal elicitation using images from International Affective Picture Session (IAPS) but there are no studies until now that have related HRV entropy with remember cognitive function.

In ET, cases where ads were unrecalled showed longer duration of visits to the brand when it appeared on screen. Higher values for this metric are related to difficulty in identifying

an object (Goldberg et al., 2002; Wedel, 2013). This fact could be explained by poor identification of the brand, which could indicate that the ad will be forgotten in the short term.

Regarding cases where participants rated the ad positively or negatively, brain activity was stronger in terms of PSD in the LIKE group than in the DISLIKE group. These results are congruent with another EEG studio based on the observation of pictures from the international affective picture system (Aftanas et al., 2004). In HRV, there were only significant differences in energy in the LF frequency band associated with sympathetic activation. In ET, no significant differences were found for the proposed metrics.

The second hypothesis was supported showing significant correlations between physiological and eye-tracking responses with the ACE score and with the number of online visits. All the indexes (pleasantness and interest) calculated showed significant correlations. The pleasantness index in the theta band presented an especially high correlation with the number of visits and the ACE score. In ET, there were relevant correlations with both metrics proposed in this work. Brand_ratio showed a positive linear relation between the percentage of time watching a brand and the number of views on the Internet. The number of fixations for the brand showed a positive relation with the number of views on the Internet. Regarding HRV, there was a positive correlation with tNN50 for both outputs. All the energy values in the frequency bands (total, LF and HF) were negatively correlated with both outputs. The ACE score was positively related to the normalized LF band associated with sympathetic activation. Non-linear entropy parameters also showed a negative correlation with both outputs revealing that an increase in the complexity of the HRV signal is associated with less quality and effectiveness of the ad on the Internet.

Hypothesis three aimed to test whether ANN with relevant biometrics could represent an interesting technique to classify ads based on their ranking on the Internet and to estimate the number of visits on the Internet. The results obtained showed that the ANN were able to accurately classify and estimate the effectiveness of each ad on the Internet via their biometric response. The results for the first network, which were adjusted to classify each ad based on a four-level ranking, showed a global average accuracy of 82.9%. Poor accuracy was obtained with ads with a lower number of views but this could have been improved if more ads in this ranking had been selected in the stimulus group. The relevant metrics for this classification were the pleasantness index and II in the theta band, the mean heart rate and the SD1 Poincare in HRV, the number of fixations during the ad and inside the brand for ET. Results for the second network to estimate the number of views on the YouTube for each ad showed a relative error of 0.199. The most important metrics for this estimate were the pleasantness index in the theta band, entropy in HRV and the number of fixations during advertising in ET. Despite good results to estimate the number of views, it would seem that classifying ads according to a ranking constitutes a better approach, taking into account the excellent results obtained in the first classifier.

Further research is needed, with more studies comparing new techniques for classification, such as Linear Discriminant Analysis, Marquardt Backpropagation Algorithm, and Deep Learning. In addition, new metrics and new signals extracted from biometric responses must be tested to find out which parameters are best to evaluate the effectiveness of advertising. The group of different categories of advertising must also be increased. Future studies should also focus on adjusting personalized classifiers to advertising categories (fashion, food, social, etc.), different channels (e.g., Facebook) and formats (desktop or mobile). Also, new metrics like facial gesture coding (McDuff et al., 2015) and fNIRS (Kopton and Kenning, 2014) could be mixed in new models for predicting Ad effectiveness.

CONCLUSION

This study has shown that aspects related to the impact of advertising, such as whether the ad is going to be remembered or whether it is going to be highly rated can be detected from an analysis of consumers' biometric responses during the viewing of these ads. We also found differences in the impact of advertising in terms of gender, which encourages the use of these biometric data to design advertising content that is tailored to each individual group of population. Other variables, such as age, cultural level, and even personality could be explored in future studies to test whether there are similar differences to the ones found in gender. The final conclusion that this study has yielded is that the effectiveness of a new ad on YouTube can be predicted using metrics extracted from EEG, HRV and ET. Up until now, there has been no evidence that biometric responses can help to classify the numbers of views on YouTube for an ad. This study has also contributed with two new metrics for ET that can be used in research on advertising. These results will help to explain the success of advertising responses showing an interesting methodology to be use by practitioners designing advertising content.

AUTHOR CONTRIBUTIONS

JG is the corresponding author. JG, EB, and MAR designed the study. JAA conducted the study. JG, EB, and JAA conducted the literature review and wrote the research summaries. ACG and VNO analyzed the EEG data, JP and JAA analyzed HRV and ET data. MAR and EB are the directors of this work. JG wrote the first draft of the manuscript, and all authors contributed to and have approved the final manuscript. Authors had full access to the study data.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01808/full#supplementary-material

REFERENCES

Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., and Suri, J. S. (2006). Heart rate variability: a review. *Med. Biol. Eng. Comput.* 44, 1031–1051. doi: 10.1007/s11517-006-0119-0

Aftanas, L. I., Reva, N. V., Varlamov, A. A., Pavlov, S. V., and Makhnev, V. P. (2004). Analysis of evoked EEG synchronization and desynchronization in conditions of emotional activation in humans: temporal and topographic characteristics. *Neurosci. Behav. Physiol.* 34, 859–867. doi: 10.1023/B:NEAB.0000038139.39812.eb

Astolfi, L., Fallani, F. D. V., Cincotti, F., Mattia, D., Bianchi, L., Marciani, M. G., et al. (2008). Neural basis for brain responses to TV commercials: a high-resolution EEG study. *IEEE Trans. Neural Syst. Rehabil. Eng.* 16, 522–531. doi: 10.1109/TNSRE.2008.2009784

Astolfi, L., Fallani, F. D. V., Cincotti, F., Mattia, D., Bianchi, L., Marciani, M. G., et al. (2009). Brain activity during the memorization of visual scenes from TV commercials: an application of high resolution EEG and steady state somatosensory evoked potentials technologies. *J. Physiol.* 103, 333–341. doi: 10.1016/j.jphysparis.2009.07.002

Baack, D. W., Wilson, R. T., and Till, B. D. (2008). Creativity and memory effects: recall, recognition, and an exploration of nontraditional media. *J. Advert.* 37, 85–94. doi: 10.2753/JOA0091-3367370407

Bellman, S., Murphy, J., Treleaven-Hassard, S., O'Farrell, J., Qiu, L., and Varan, D. (2013). Using internet behavior to deliver relevant television commercials. *J. Interact. Mark.* 27, 130–140. doi: 10.1016/j.intmar.2012.12.001

Bigné, E. (2016). Frontiers in research in business: will you be in? *Eur. J. Manag. Bus. Econ.* 25, 89–90. doi: 10.1016/j.redeen.2016.09.001

Bigné, E., Llinares, C., and Torrecilla, C. (2016). Elapsed time on first buying triggers brand choices within a category: a virtual reality-based study. *J. Bus. Res.* 69, 1423–1427. doi: 10.1016/j.jbusres.2015.10.119

Blanco-Velasco, M., Weng, B., and Barner, K. E. (2008). ECG signal denoising and baseline wander correction based on the empirical mode decomposition. *Comput. Biol. Med.* 38, 1–13. doi: 10.1016/j.compbiomed.2007.06.003

Boksem, M. A. S., and Smidts, A. (2014). Brain responses to movie-trailers predict individual preferences for movies and their population-wide commercial success. *J. Mark. Res.* 52, 482–492. doi: 10.1509/jmr.13.0572

Bolls, P. D., Wise, K., and Bradley, S. D. (2012). "Embodied motivated cognition: a theoretical framework for studying dynamic mental processes underlying advertising exposure," in *Advertising Theory*, eds S. Rodgers and E. Thorson (New York, NY: Routledge), 105–119.

Bradley, M. M., Houbova, P., Miccoli, L., Costa, V. D., and Lang, P. J. (2011). Scan patterns when viewing natural scenes: emotion, complexity, and repetition. *Psychophysiology* 48, 1544–1553. doi: 10.1111/j.1469-8986.2011.01223.x

Castiglioni, P., and Di Rienzo, M. (1996). "On the evaluation of heart rate spectra: the lomb periodogram," in *Proceedings of the Computers in Cardiology, IEEE*, Indianapolis, IN, 505–508. doi: 10.1109/CIC.1996.542584

Cha, M., Kwak, H., Rodriguez, P., Ahn, Y. Y., and Moon, S. (2007). "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement* (San Diego, CA: Association for Computing Machinery), 1–14. doi: 10.1145/1298306.1298309

Chen, L., Zhou, Y., and Chiu, D. M. (2014). "A lifetime model of online video popularity," in *Proceedings of the 23rd International Conference on Computer Communication and Networks, IEEE*, Shanghai, 1–8. doi: 10.1109/ICCCN.2014.6911774

Christoforou, C., Christou-Champi, S., Constantinidou, F., and Theodorou, M. (2015). From the eyes and the heart: a novel eye-gaze metric that predicts video preferences of a large audience. *Front. Psychol.* 6:579. doi: 10.3389/fpsyg.2015.00579

Clifford, G. D. (2006). "ECG statistics, noise, artifacts, and missing data," in *Advanced Methods and Tools for ECG Data Analysis*, eds G. Cliford, F. Azuaje, and P. E. MsSahrry (Boston, MA: Artech House), 55–99.

Colomer Granero, A., Fuentes-Hurtado, F., Naranjo, V., Guixeres, J., Ausín, J. M., and Alcañiz, M. (2016). A comparison of physiological signal analysis techniques and classifiers for automatic emotional evaluation of audiovisual contents". *Front. Comput. Neurosci.* 10:74. doi: 10.3389/fncom.2016.00074

Couwenberg, L. E., Boksem, M. A., Dietvorst, R. C., Worm, L., Verbeke, W. J., and Smidts, A. (2017). Neural responses to functional and experiential ad appeals: explaining ad effectiveness. *Int. J. Res. Mark.* 34, 355–366. doi: 10.1016/j.ijresmar.2016.10.005

Curry, B., and Moutinho, L. (1993). Neural networks in marketing: modelling consumer responses to advertising stimuli. *Eur. J. Mark.* 27, 5–20. doi: 10.1108/03090569310040325

Daugherty, T., Hoffman, E., and Kennedy, K. (2016). Research in reverse: Ad testing using an inductive consumer neuroscience approach. *J. Bus. Res.* 69, 3168–3176. doi: 10.1016/j.jbusres.2015.12.005

Davidson, R. J. (2004). What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research. *Biol. Psychol.* 67, 219–233. doi: 10.1016/j.biopsycho.2004.03.008

Deitz, G. D., Royne, M. B., Peasley, M. C., and Huang, J. C. (2016). EEG-based measures versus panel ratings: predicting social-media based behavioral responses to super Bowl Ads. *J. Advert. Res.* 56, 217–227. doi: 10.2501/JAR-2016-030

Demarzo, M. M. P., Montero-Marin, J., Stein, P. K., Cebolla, A., Provinciale, J. G., and García-Campayo, J. (2014). Mindfulness may both moderate and mediate the effect of physical fitness on cardiovascular responses to stress: a speculative hypothesis. *Front. Physiol.* 5:105. doi: 10.3389/fphys.2014.00105

Donthu, N., Cherian, J., and Bhargava, M. (1993). Factors influencing recall of outdoor advertising. *J. Advert. Res.* 33, 64–73.

dos Santos, R. D. O. J., de Oliveira, J. H. C., Rocha, J. B., and Giraldi, J. D. M. E. (2015). Eye tracking in neuromarketing: a research agenda for marketing studies. *Int. J. Psychol. Stud.* 7, 32–42. doi: 10.5539/ijps.v7n1p32

Elsen, M., Pieters, P., and Wedel, M. (2016). Thin slice impressions: how advertising evaluation depends on exposure duration. *J. Mark. Res.* 53, 563–579. doi: 10.1509/jmr.13.0398

Feise, R. J. (2002). Do multiple outcome measures require P-value adjustment? *BMC Med. Res. Methodol.* 2:8.

Fishman, M., Jacono, F. J., Park, S., Jamasebi, R., Thungtong, A., Loparo, K. A., et al. (2012). A method for analyzing temporal patterns of variability of a time series from Poincare plots. *J. Appl. Physiol.* 113, 297–306. doi: 10.1152/japplphysiol.01377.2010

Fjorback, L. O., Arendt, M., Ørnbøl, E., Fink, P., and Walach, H. (2011). Mindfulness-based stress reduction and mindfulness-based cognitive therapy–a systematic review of randomized controlled trials. *Acta Psychiatr. Scand.* 124, 102–119. doi: 10.1111/j.1600-0447.2011.01704.x

Gao, J. F., Yang, Y., Lin, P., Wang, P., and Zheng, C. X. (2010). Automatic removal of eye-movement and blink artifacts from EEG signals. *Brain Topogr.* 23, 105–114. doi: 10.1007/s10548-009-0131-4

Geisler, F. C., Vennewald, N., Kubiak, T., and Weber, H. (2010). The impact of heart rate variability on subjective well-being is mediated by emotion regulation. *Pers. Individ. Dif.* 49, 723–728. doi: 10.1016/j.paid.2010.06.015

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., and Wichansky, A. M. (2002). "Eye tracking in web search tasks," in *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications - ETRA*, New York, NY, 51. doi: 10.1145/507072.507082

Grandjean, D., Sander, D., and Scherer, K. R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Conscious. Cogn.* 17, 484–495. doi: 10.1016/j.concog.2008.03.019

Guerreiro, J., Rita, P., and Trigueiros, D. (2015). Attention, emotions and cause-related marketing effectiveness. *Eur. J. Mark.* 49, 1728–1750. doi: 10.1108/EJM-09-2014-0543

Guixeres, J., Redon, P., Saiz, J., Alvarez, J., Torro, M. I., Cantero, L., et al. (2014). Cardiovascular fitness in youth: association with obesity and metabolic abnormalities. *Nutr. Hosp.* 29, 1290–1297. doi: 10.3305/nh.2014.29.6.7383

Ha, L. (2008). Online advertising research in advertising journals: a review. *J. Curr. Issues Res. Advert.* 30, 31–48. doi: 10.1080/10641734.2008.10505236

Harmon-Jones, E., Gable, P. A., and Peterson, C. K. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: a review and update. *Biol. Psychol.* 84, 451–462. doi: 10.1016/j.biopsycho.2009.08.010

Holmqvist, K., Andrà, C., Lindström, P., Arzarello, F., Ferrara, F., Robutti, O., et al. (2011). A method for quantifying focused versus overview behavior in AOI sequences. *Behav. Res. Methods* 43, 987–998. doi: 10.3758/s13428-011-0104-x

Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 371–375.

Kent, R. J., and Allen, C. T. (1994). Competitive interference effects in consumer memory for advertising: the role of brand familiarity. *J. Mark.* 58, 97–105. doi: 10.2307/1252313

Khushaba, R. N., Wise, C., Kodagoda, S., Louviere, J., Kahn, B. E., and Townsend, C. (2013). Consumer neuroscience: assessing the brain response to marketing stimuli using electroencephalogram (EEG) and eye tracking. *Expert Syst. Appl.* 40, 3803–3812. doi: 10.1016/j.eswa.2012.12.095

Kim, K., Hayes, J. L., Avant, J. A., and Reid, L. N. (2014). Trends in advertising research: a longitudinal analysis of leading advertising, marketing, and communication journals, 1980 to 2010. *J. Advert.* 43, 296–316. doi: 10.1080/00913367.2013.857620

Kopton, I. M., and Kenning, P. (2014). Near-infrared spectroscopy (NIRS) as a new tool for neuroeconomic research. *Front. Hum. Neurosci.* 8:549. doi: 10.3389/fnhum.2014.00549

Kühn, S., Strelow, E., and Gallinat, J. (2016). Multiple "buy buttons" in the brain: forecasting chocolate sales at point-of-sale based on functional brain activation using fMRI. *Neuroimage* 136, 122–128. doi: 10.1016/j.neuroimage.2016.05.021

Lang, A., Bolls, P., Potter, R. F., and Kawahara, K. (1999). The effects of production pacing and arousing content on the information processing of television. *J. Broadcast. Electron. Media* 43, 451–475. doi: 10.1080/08838159909364504

Lee, J., and Ahn, J. H. (2012). Attention to banner ads and their effectiveness: an eye-tracking approach. *Int. J. Electron. Commerce* 17, 119–137. doi: 10.2753/JEC1086-4415170105

McAlister, L., Srinivasan, R., Jindal, N., and Cannella, A. A. (2016). Advertising effectiveness: the moderating effect of firm strategy. *J. Mark. Res.* 53, 207–224. doi: 10.1509/jmr.13.0285

McDuff, D., El Kaliouby, R., Cohn, J. F., and Picard, R. W. (2015). Predicting ad liking and purchase intent: large-scale analysis of facial responses to ads. *IEEE Trans. Affect. Comput.* 6, 223–235. doi: 10.1109/TAFFC.2014.2384198

McKinsey (2015). *Global Media Report 2015.* Available at: www.mckinsey.com/client_service/media_and_entertainment (accessed December 6, 2016).

Mognon, A., Jovicich, J., Bruzzone, L., and Buiatti, M. (2011). ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology* 48, 229–240. doi: 10.1111/j.1469-8986.2010.01061.x

Moriarty, S., Mitchell, N., and Wells, W. (2012). *Advertising & IMC Principles and Practice*, 9th Edn. Hoboken, NJ: Pearson.

Morin, C. (2011). Neuromarketing: the new science of consumer behavior. *Society* 48, 131–135. doi: 10.1109/MPUL.2012.2189166

Mould, D., Mandryk, R. L., and Li, H. (2012). Emotional response and visual attention to non-photorealistic images. *Comput. Graph.* 36, 658–672. doi: 10.1016/j.cag.2012.03.039

Ohme, R., Reykowska, D., Wiener, D., and Choromanska, A. (2010). Application of frontal EEG asymmetry to advertising research. *J. Econ. Psychol.* 31, 785–793. doi: 10.3791/55872

Pan, J., and Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* 32, 230–236. doi: 10.1109/TBME.1985.325532

Pieters, R., Warlop, L., and Wedel, M. (2002). Breaking through the clutter: benefits of advertisement originality and familiarity for brand attention and memory. *Manage. Sci.* 48, 765–781. doi: 10.1287/mnsc.48.6.765.192

Piskorski, J., and Guzik, P. (2007). Geometry of the Poincaré plot of RR intervals and its asymmetry in healthy adults. *Physiol. Meas.* 28, 287–300. doi: 10.1088/0967-3334/28/3/005

Richman, J. S., and Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* 278, H2039–H2049.

Rossiter, J. R., and Eagleson, G. (1994). Conclusions from the ARF's copy research validity project. *J. Advert. Res.* 34, 19–33.

Ruanguttamanun, C. (2014). Neuromarketing: I put myself into a fMRI scanner and realized that I love Louis Vuitton Ads. *Procedia* 148, 211–218. doi: 10.1016/j.sbspro.2014.07.036

Shehu, E., Bijmolt, T. H., and Clement, M. (2016). Effects of likeability dynamics on consumers' intention to share online video advertisements. *J. Interact. Mark.* 35, 27–43. doi: 10.1016/j.intmar.2016.01.001

Smith, A. N., Fischer, E., and Yongjian, C. (2012). How does brand-related user-generated content differ across YouTube, Facebook, and Twitter? *J. Interact. Mark.* 26, 102–113. doi: 10.1016/j.intmar.2012.01.002

Strach, P., Zuber, K., Fowler, E. F., Ridout, T. N., and Searles, K. (2015). In a different voice? Explaining the use of men and women as voice-over announcers in political advertising. *Polit. Commun.* 32, 183–205. doi: 10.1080/10584609.2014.914614

Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology (1996). Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Eur. Heart J.* 17, 354–381. doi: 10.1093/oxfordjournals.eurheartj.a014868

Tomkovick, C., Yelkur, R., and Christians, L. (2001). The USA's biggest marketing event keeps getting bigger: an in-depth look at Super Bowl advertising in the 1990s. *J. Mark. Commun.* 7, 89–108. doi: 10.1080/13527260121725

Vakratsas, D., and Ambler, T. (1999). How advertising works: What do we really know? *J. Mark.* 63, 26–43. doi: 10.2307/1251999

Valenza, G., Allegrini, P., Lanatà, A., and Scilingo, E. P. (2012). Dominant Lyapunov exponent and approximate entropy in heart rate variability during emotional visual elicitation. *Front. Neuroeng.* 5:3. doi: 10.3389/fneng.2012.00003

Valenza, G., Citi, L., Lanatá, A., Scilingo, E. P., and Barbieri, R. (2014). Revealing real-time emotional responses: a personalized assessment based on heartbeat dynamics. *Sci. Rep.* 4:4998. doi: 10.1038/srep04998

Varan, D., Lang, A., Barwise, P., Weber, R., and Bellman, S. (2015). How reliable are neuromarketers' measures of advertising effectiveness? *J. Advert. Res.* 55, 176–191. doi: 10.2501/JAR-55-2-176-191

Vecchiato, G., Astolfi, L., Tabarrini, A., Salinari, S., Mattia, D., Cincotti, F., et al. (2010). EEG analysis of the brain activity during the observation of commercial, political, or public service announcements. *Comput. Intell. Neurosci.* 2010:985867. doi: 10.1155/2010/985867

Vecchiato, G., Maglione, A. G., Cherubino, P., Wasikowska, B., Wawrzyniak, A., Latuszynska, A., et al. (2014). Neurophysiological tools to investigate consumer's gender differences during the observation of TV commercials. *Comput. Math. Methods Med.* 2014:912981. doi: 10.1155/2014/912981

Vecchiato, G., Susac, A., Margeti, S., Fallani, F. D. V., Maglione, A. G., Supek, S., et al. (2013). High-resolution EEG analysis of power spectral density maps and coherence networks in a proportional reasoning task. *Brain Topogr.* 26, 303–314. doi: 10.1007/s10548-012-0259-5

Vecchiato, G., Toppi, J., Astolfi, L., Fallani, F. D. V., Cincotti, F., Mattia, D., et al. (2011). Spectral EEG frontal asymmetries correlate with the experienced pleasantness of TV commercial advertisements. *Med. Biol. Eng. Comput.* 49, 579–583. doi: 10.1007/s11517-011-0747-x

Venkatraman, V., Dimoka, A., Pavlou, P. A., Vo, K., Hampton, W., Bollinger, B., et al. (2015). Predicting advertising success beyond traditional measures: new insights from neurophysiological methods and market response modeling. *J. Mark. Res.* 52, 436–452. doi: 10.1509/jmr.13.0593

Verhellen, Y., Dens, N., and De Pelsmacker, P. (2013). Consumer responses to brands placed in youtube movies: the effect of prominence and endorser expertise. *J. Electron. Commerce Res.* 14, 287–303.

Wackermann, J., Lehmann, D., Michel, C. M., and Strik, W. K. (1993). Adaptive segmentation of spontaneous EEG map series into spatially defined microstates. *Int. J. Psychophysiol.* 14, 269–283. doi: 10.1016/0167-8760(93)90041-M

Wedel, M. (2013). *Attention Research in Marketing: A Review of Eye Tracking Studies.* 1–28. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2460289

Wedel, M., and Kannan, P. K. (2016). Marketing analytics for data-rich environments. *J. Mark.* 80, 97–121. doi: 10.1509/jm.15.0413

Wedel, M., and Pieters, R. (2014). "Looking at vision: eye/face/head tracking of consumers for improved marketing decisions," in *The Routledge Companion to the Future of Marketing*, eds L. Moutinho, E. Bigné, and A. K. Manrai (New York, NY: Routledge), 177–189.

Werkle-Bergner, M., Müller, V., Li, S. C., and Lindenberger, U. (2006). Cortical EEG correlates of successful memory encoding: implications for lifespan comparisons. *Neurosci. Biobehav. Rev.* 30, 839-854. doi: 10.1016/j.neubiorev.2006.06.009

West, P. M., Brockett, P. L., and Golden, L. L. (1997). A comparative analysis of neural networks and statistical methods for predicting consumer choice. *Mark. Sci.* 16, 370-391. doi: 10.1287/mksc.16.4.370

Zhou, R., Khemmarat, S., Gao, L., Wan, J., and Zhang, J. (2016). How YouTube videos are discovered and its impact on video views. *Multimed. Tools Appl.* 75, 6035-6058. doi: 10.1007/s11042-015-3206-0