International Journal of Advertising
The Review of Marketing Communications

ISSN: 0265-0487 (Print) 1759-3948 (Online) Journal homepage: www.tandfonline.com/journals/rina20

How to accurately measure attention to video advertising

Nicole Hartnett, Steven Bellman, Virginia Beal, Rachel Kennedy, Claire Charron & Duane Varan

INTERNATIONAL JOURNAL OF ADVERTISING
2025, VOL. 44, NO. 1, 184–207
https://doi.org/10.1080/02650487.2024.2435164

How to accurately measure attention to video advertising

Nicole Hartnett[a], Steven Bellman[a], Virginia Beal[a], Rachel Kennedy[a], Claire Charron[b] and Duane Varan[c]

[a]Ehrenberg-Bass Institute for Marketing Science, University of South Australia, Adelaide, SA, Australia; [b]Market Insights, Google, London, UK; [c]MediaScience, Austin, TX, USA

ABSTRACT

Attention has been proposed as a replacement for ratings to better measure advertising exposure quality. Attention measurement suppliers, who are key stakeholders in this proposed shift, have primarily used eye tracking to assess how people pay attention to ads. However, other literature-supported physiological measures may be beneficial or superior for measuring attention to advertising (e.g. heart rate, skin conductance). This research used a lab-based experiment that manipulated attention levels to video ads (i.e. high vs low) to test the accuracy of a range of potentially scalable physiological attention measures. Eight physiological measures were collected and compared to electroencephalogram (EEG), a 'gold standard' but non-scalable direct measure of attention, and a survey-based self-reported attention measure. The results suggest that eye tracking measures have problems discriminating high attention from low attention. Heart rate proved best at this discrimination task and has the advantage of being able to measure attention to sound as well as visuals.

ARTICLE HISTORY
Received 29 September 2023
Accepted 20 November 2024

KEYWORDS
Attention; advertising; video; eye tracking; heart rate; EEG

Introduction

Marketers have increasingly become interested in attention to advertising. As the number of media platforms vying for attention rises (Kim, Kang, and Lee 2021), coupled with claims that attention spans have shortened, the result is an advertising-saturated environment where brands compete aggressively for limited customer attention. With proliferating advertising options, another practical concern for advertisers is the incomparability of measures across media, which have differing definitions of 'viewable impressions', making it difficult to assess the relative value of advertising placements. Responding to these challenges, some agencies and associations (Dentsu 2021; The Attention Council 2021) have championed attention as the

CONTACT Nicole Hartnett Nicole.Hartnett@marketingscience.info Ehrenberg-Bass Institute for Marketing Science, University of South Australia, Adelaide, SA, Australia
This article has been corrected with minor changes. These changes do not impact the academic content of the article.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/02650487.2024.2435164.

new media currency to compare and price media/format options. A visual attention measure (specifically eyes on screen) is the basis for a proposed prototype media buying pricing structure using attention-based cost per mille (CPM): 'cost per 1,000 attentive seconds' (The Attention Council 2021).

Yet, attention is not a single construct with a single measure, and our understanding of attention is still far from complete (Krauzlis et al. 2023; Oberauer 2019; Ramsøy 2019; Rosenholtz 2024). Various sources agree that attention is the controlled, but also the automatic, focusing of limited cognitive resources on external stimuli, via any of the senses, or on internal thoughts invisible to devices like webcams (Cowan 1995; James 1890; Shiffrin and Schneider 1977; Venkatraman et al. 2015). For this reason, this research investigates and systematically compares the validity of eight promising physiological measures that potentially measure both external and internal attention, by comparing them with electroencephalogram (EEG) measurement of brainwaves, which is considered a 'gold standard' but unscalable direct measure of attention (Bazzani et al. 2020; Kolar et al. 2021; Reeves et al. 1985; Simons et al. 2003). Priority is given to measures most likely to be scalable in field, meaning they can be collected with large samples at comparatively low costs, as this will be critical for any industry currency.

The approach was to identify a ground truth sample of 10 video ads selected from hundreds of contenders through extensive pre-testing. Pre-tests identified five ads as high attention and another five as low attention. In a lab, participants watched programs including these ads while their physiological responses were monitored continuously. Using literature-derived 'signatures' (translated into hypotheses) for each measure, we tested how consistently the measures discriminated high from low attention ads.

The results contribute new knowledge about attention by identifying literature-derived signatures that distinguish low and high attention for eight promising scalable physiological measures. These signatures were evidence of orienting responses compared with a person's resting baseline. We compared responses on these measures to a sample of typical ads. Results indicate that only one of the measures, heart rate, validly measures attention in line with its hypothesized signature. We discuss these results using dual processing theories, which correlate high attention with greater message involvement and cognitive processing, while at the same time explaining how ads can be effective under conditions of low attention and involvement (Greenwald and Leavitt 1984; Krugman 1965; MacInnis and Jaworski 1989; Petty, Cacioppo, and Schumann 1983; Santoso et al. 2022). For researchers, a methodological contribution is providing insights into how to implement physiological measures of attention. For practitioners, these results contribute by showing the relative performance of a range of scalable attention measures, as potential candidates for measuring attention, including as a new currency for media buying.

Research background and hypotheses

Models or frameworks of attention often consider levels of attention (Graham 1997; Greenwald and Leavitt 1984; Krugman 1965; Oberauer 2019; Petty, Cacioppo, and Schumann 1983). Common classifications include 'low' and 'high', 'passive' and 'active', 'divided' and 'focused', 'partial' and 'full', or 'automatic' and 'controlled' attention. Some incorporate preattention (Greenwald and Leavitt 1984; Heath 2000; Janiszewski 1993), also called inattention, which is an automatic environment-scanning process that uses very little attentional capacity. We follow prior research in distinguishing between two levels of attention, specifically low versus high attention (MacInnis and Jaworski 1989).

Low attention occurs when the amount of attention paid to the external world, here specifically to advertising, is no higher than a person's resting baseline level. Attention could be even lower than this baseline because the person is asleep (Regan, Hallett, and Gordon 2011). But when awake, low attention is not consciously noticing specific content, but automatically monitoring sensory information while looking or listening to the world (Cowan 1995; Mack and Rock 1998; Rees et al. 1999; Sokolov and Cacioppo 1997). Despite lacking or with limited conscious processing, some of this information can be available later, as demonstrated by improved memory (Dehaene 2014).

High attention occurs when something is consciously noticed and recorded (Heath 2000), typically accompanied by moderate to high levels of cognitive processing (MacInnis and Jaworski 1989). Intuitively, more attention should be better than less but diminishing returns have been noted, whereby doubling attention to advertising does not double its effectiveness (Varan et al. 2020). Most advertisers, however, do not want to risk their advertising going unnoticed, hence the use of minimum viewability standards, such as two seconds of exposure to video ads (Media Rating Council 2019).

However, even fleeting or low attention can be sufficient to promote recognition or positive attitudes (Santoso et al. 2022; Zajonc 1968), especially for consumers who are familiar with or already buy the brand (Alba and Hutchinson 1987; Simmonds et al. 2020a). Dual process models of communication effects assume a high correlation between levels of attention and cognitive processing (Krugman 1965; MacInnis and Jaworski 1989; Santoso et al. 2022). Under low attention/involvement conditions, these models suggest advertising works via low levels of cognitive processing requiring 'easier' heuristic or peripheral processing (Chaiken 1980; Petty and Cacioppo 1986). Some argue that low attention is better for video advertising (Heath 2000, 2009), because low attention viewers generate little counterarguing (Heath, Brandt, and Nairn 2006; Petty, Wells, and Brock 1976). Even advertisers who want low attention exposure for their ads require measures that delineate between low (bordering on inattention) and high attention, which is the basis for our hypotheses.

Hypotheses for physiological measures

As outlined above, we adopt the dichotomy of low versus high attention to test the validity of physiological measures of attention to video advertising. Beyond alignment with prior literature, this choice was also partly driven by the ability to identify corresponding 'signatures' for each of the physiological measures. The term 'signature,' when used in sciences like spectroscopy (Oelofse et al. 2010), means identifying the presence of something from a pattern in the data. In this research, a sign of attention is hypothesized to be a specific pattern in physiological data. The boundary between low and high attention should be marked by any sign of attention, such as an orienting response (Greenwald and Leavitt 1984). Orienting responses are automatic, reflexive, attentional responses to changes in the environment (novel stimuli) or to stimuli that signal important information (signal stimuli) (Lang 1990; Lang et al. 2002). These can be measured using a variety of physiological measures (Potter and Bolls 2012). An orienting response indicates the presence of some controlled cognitive processing, which is characteristic of high attention and involvement as opposed to low attention automatic processing (Oberauer 2019). If orienting responses are absent when advertising is present, it is not being consciously attended to. This means no substantive cognitive resources are being allocated to processing the advertising, so it will have less of an effect on the person's memory or future behavior (Krauzlis et al. 2023).

Eight promising measures of attention to advertising were identified from prior research, along with each measure's hypothesized signature for distinguishing high from low attention. Each signature, or hypothesized pattern, refers to a change from the person's resting baseline, which is a standard procedure with biological measures to control for differences between individuals (Wainer 1991). In theory and typical measurement, the resting baseline is a common state to which people return, so it should be identifiable by all measures. The eight promising and potentially scalable measures and their associated attention-signature hypotheses are described below, as well as the 'gold standard' EEG measure to which they are compared.

- **Heart rate** is captured as ***beats per minute (BPM)***, which is the measure most often used in marketing research incorporating cardiac activity (Kakaria et al. 2023). BPM measures attention when it slows down, relative to baseline, as the body temporarily 'quiets' to intake external information (Graham and Clifton 1966; Lacey 1967; Lang 1994).

H1: *BPM less than baseline indicates high attention.*

- **Eye tracking** observes what people look at, for how long, and how often. ***Eyes on screen (EOS)*** indicates that a person is present and looking at the screen (Thorson 1994). In this research, whether a webcam detected eyes looking at the screen (i.e., EOS = 1 vs 0 otherwise) was sampled 30 times a second. Average EOS measured the percentage of the ad's duration in which the person is looking at the ad. High-attention advertising should be associated with participants looking at the screen more than during a meditative video designed to establish a resting baseline.

H2: *EOS greater than baseline indicates high attention.*

- ***Blink duration*** is an alternative measure of visual attention, which has been used to detect driver inattention (eyes closed). Longer blink durations suggest increasing fatigue and lower attention (Morris and Miller 1996; Stern, Boyer, and Schroeder 1994). The inverse implies that shorter blink durations, relative to baseline, are a signature of high attention.

H3: *Blink duration less than baseline indicates high attention.*

- More sophisticated eye tracking measures identify when the eyes pause or fixate on a specific area. A high number of ***fixations per second (FPS)*** in an area of interest (AOI), the ad, relative to baseline, indicates information processing (Gordon-Hecker et al. 2020; Heath, Nairn, and Bottomley 2009; Rayner 1998; Simmonds et al. 2020a), consistent with higher attention. FPS, therefore, potentially measures not only automatic visual attention but also controlled ad processing, and so is arguably a better measure of attention than EOS 'dwell time' or fixation duration (Gidlöf et al. 2013; Orquin, Bagger, and Loose 2013).

H4: *FPS greater than baseline indicates high attention.*

- Fixation location can also be used to measure differences in where individuals look. Video directors encourage audiences to look at the same central location on-screen (Loschky et al. 2015) so high ***individual attention dispersion (IAD),*** when an individual's fixations disperse away from where others are looking, indicates disinterest, and has been shown to predict advertising avoidance (Teixeira, Wedel, and Pieters 2010). This finding suggests that low dispersion represents concentrated, and higher, attention.

H5: *IAD less than baseline indicates high attention.*

- **Skin conductance** (SC) measures changes in the conductance of electricity due to changes in the moisture on the skin, caused by the body's response to heat, and by increases in sympathetic nervous system (SNS) stimulation, especially for eccrine sweat glands on the hands and feet (Benedek and Kaernbach 2010). When attentional reflexes (or orienting responses) to emotional or signal stimuli cause spikes in SNS arousal, this increases sweating and suggests SC is a potential measure of attention, as well as the best measure of SNS arousal (Potter and Bolls 2012). We measured dynamic or short-term SC activity (***phasic SC***), reflecting these short-term orienting responses, averaged over the duration of an ad, and longer-lasting step-changes in average SNS arousal (change in ***SC level or SCL***, also averaged over ad duration). If more arousing stimuli receive greater attention, phasic SC and SCL are both potentially useful measures of attention and ad processing, in the form of high arousal emotional responses (Koruth et al. 2015).

H6: *Phasic SC greater than baseline indicates high attention.*

H7: *SCL greater than baseline indicates high attention.*

- **Facial expression** reflects emotional response (SNS activity), which is related to attention. Recent research has cast some doubt on whether facial expressions are valid indicators of felt emotions (Barrett 2022; Barrett et al. 2019) but smiling in the context of ad viewing has been associated with effectiveness (McDuff et al. 2015), commonly via humour appeals (Bellman et al. 2017; Hartnett et al. 2016). The facial expression has traditionally been human-coded (Ekman and Friesen 1978) or measured using facial electrodes, called facial electromyography (fEMG) (Lajante, Droulers, and Amarantini 2017). fEMG is the most accurate facial measure of attention, but it requires skilled electrode placement (Varan et al. 2015) as well as the right sampling rate and filtering (Lajante, Droulers, and Amarantini 2017), and for these reasons is not easily scalable. Large-scale measurement of facial expression requires computer coding of visible, conscious expressions, recorded by webcams (McDuff et al. 2015). Facial coding from webcams can underestimate the amount of attention video ads receive compared with best practice fEMG measurement (Lajante, Droulers, and Amarantini 2017), but there is evidence that computer coding is a valid alternative to measuring 'joy,' albeit with time delays (Westermann et al. 2024). Further, visible smiling, detectable by a webcam, is unlikely to overstate attention because 'getting a joke' normally requires attention with consciousness (Koch and Tsuchiya 2007) and extensive cognitive processing (Coulson and Kutas 2001; MacInnis and Jaworski 1989). For this reason, we hypothesize that if the average probability of smiling during the ad is higher than the baseline, that is a potential sign of high attention.

H8: *Smiling greater than baseline indicates high attention.*

Ideally, these measures should be compared with a 'gold standard' direct measure of attention (and ad processing), such as brain imaging, including functional magnetic resonance imaging (fMRI) (Couwenberg et al. 2017; Sung et al. 2019; Venkatraman et al. 2015). However, fMRI is not scalable (Weber, Mangus, and Huskey 2015). Instead, this study used electroencephalogram (EEG) as the direct measure of attention activity in the brain for its 'gold standard' comparison (Bazzani et al. 2020; Kolar et al. 2021; Reeves et al. 1985; Simons et al. 2003) with additional triangulation between the eight physiological measures and a self-reported attention measure, given known issues relying on a single measure (Kolar et al. 2021). EEG is typically not commercially scalable because it requires measurement in labs but was sufficiently scalable and affordable for the purposes of this study. Lower frequency ***EEG alpha*** brain waves indicate low attention; therefore, an EEG signature of attention is alpha dipping below the baseline level (Kolar et al. 2021; Reeves et al. 1985).

H9: *EEG Alpha less than baseline indicates high attention.*

Table 1 lists the attention-signature hypotheses, along with the key studies justifying each hypothesis.

Method

This study used three stages to select and test differences between ads expected to be high versus low in attention-getting capacity.

Advertising selection

Identifying the sample of ads started by sourcing 324 video ads that were all 30-seconds long and had previously aired in the USA. These ads were collated from

Table 1. Attention signature hypotheses and supporting literature.

Heart Rate	Eye Tracking				Skin Conductance		Facial Coding	EEG
BPM	EOS	Blink Duration	FPS	IAD	Phasic SC	SCL	Smiling	Alpha
A ***negative*** change from baseline in BPM will indicate high attention*	A ***positive*** change from baseline in EOS will indicate high attention	A ***negative*** change from baseline in blink duration will indicate high attention	A ***positive*** change from baseline in FPS will indicate high attention	A ***negative*** change from baseline in IAD will indicate high attention	A ***positive*** change from baseline in the phasic driver component of SC will indicate high attention	A ***positive*** change from baseline in SCL will indicate high attention	A ***positive*** change from baseline in smiling activity will indicate high attention	A ***negative*** change from baseline in EEG alpha wave activity will indicate high attention
H1 **<B**	**H2** **>B**	**H3** **<B**	**H4** **>B**	**H5** **<B**	**H6** **>B**	**H7** **>B**	**H8** **>B**	**H9** **<B**
(Graham and Clifton 1966; Greenwald and Leavitt 1984; Koruth et al. 2015; Lacey 1967; Lang 1994; Öhman 1997; Potter and Bolls 2012)	(Segijn et al. 2017; Thorson 1994)	(Morris and Miller 1996; Stern et al. 1994; Tijerina et al. 1999)	(Heath et al. 2009; Rayner 1998; Simmonds et al. 2020a)	(Loschky et al. 2015; Teixeira et al. 2010)	(Benedek and Kaernbach 2010; Bolls et al. 2001; Potter and Bolls 2012)	(Lee and Lang 2009; MacInnis et al. 1991)	(McDuff et al. 2015; Russell and Barrett 1999)	(Boksem and Smidts 2015; Gevins et al. 1979; Reeves et al. 1985; Venkatraman et al. 2015)

*High attention as opposed to low attention (the null hypothesis).
BPM: beats per minute; EOS: eyes on screen; FPS: fixations per second; IAD: individual attention dispersion; Phasic SC: phasic driver component of skin conductance; SCL: skin conductance level; EEG: electroencephalogram; B = (resting) baseline.

a consumer goods manufacturer, a technology company, and publicly accessible databases. Using expert judgement, the authors selected 100 promising ads including diverse creative characteristics and a mix of brands and categories, which were then rated by consumers in an online survey. These 100 ads were chosen for consumer testing in anticipation of their content eliciting varied attention levels, drawing on relevant literature (e.g. Bellman et al. 2019; Wooley et al. 2022). Based on the results of that survey, 10 ads were selected for the final lab study to best manipulate high versus low attention ads.

Online pre-testing survey

A consumer pre-test to measure the attention-getting capacity of 100 ads was conducted via an online survey with a nationally representative sample of 636 consumers living in the USA, who were recruited and rewarded by Toluna, a large market research panel provider. Each respondent saw 10 ads randomly allocated from the 100 ads, presented in a random order in a continuous reel lasting 5-minutes. Respondents were instructed to watch normally, like they would watch content on a TV or another device at home. Afterwards, participants completed a survey where they were presented with images of three key scenes from the 10 ads they had seen, one ad at a time, and asked to rate how much attention they had paid on a 0% to 100% attention scale to each ad; where 0% means they gave practically no attention throughout the ad, while 100% means they gave all their attention throughout the ad. This self-reported attention scale has been validated with manipulations of behavioral multitasking (Chang and Thorson 2023; Segijn et al. 2017). Recent theoretical developments also support asking respondents as the most direct and reliable way to measure their state of attention to sensory stimuli (Krauzlis et al. 2023). Each respondent gave 10 ratings, with each ad rated by at least 40 respondents.

Selection of test ads

From the survey responses, the authors identified 10 ads to include in the laboratory study that represent the relative extremes of attention levels. Five ads that rated among the lowest on the attention scale (mean scores from 56% to 62%) were classified for the experiment as low attention (i.e. they should show no more attention than resting baseline), and five ads that rated among the highest (74% to 82%) were classified as high attention (i.e. they should get more attention than resting baseline). We acknowledge that, objectively, our 'low' ads were rated as 'medium' attention (i.e. just over 50% on the attention scale), and this may have presented problems for testing our resting-baseline-comparison hypotheses. However, they were the lowest attention-getting ads from the set of typical ads, and ANOVA tests of the 10 final ads identified two groups of ads that differed significantly between groups but did not differ within groups ($F(9, 609) = 6.12$, $p < 0.001$; Tukey B homogenous subset 1 = low_1 to low_5, vs subset 4 = high_1 to high_5; the lowest scoring ad [74%] in the high attention group had significantly higher attention than the highest scoring ad [62%] in the low attention group: $t(117) = -1.82$, $p = 0.04$ [1-tailed]).

The ads were from diverse categories, specifically financial services, Internet service providers, online retail, skin care, automotive, whitegoods, electronics, snacks, side dishes, and household cleaners, and from 10 different brands (allowing manipulation checks of brand recall). The ads varied in their creative execution, such as using emotional or rational appeals, demonstrating products or services, showing people or animals, reliance on voice overs, amount of branding, and so forth. The attention manipulation has face validity in that the high attention ads contain more attention-getting elements identified in prior research compared with the low attention ads (e.g. animals).

Laboratory study

Procedure

For the laboratory experiment, 261 participants (of whom 181 yielded complete data across all dependent variables, including the survey variables; the analyses used all the data available for each measure) sat at individual desks with a large computer monitor, acting like a flat-screen television. Participants were fitted with electrodes and completed exercises to calibrate the eye tracking and facial coding devices. Their physiological responses were then continuously measured throughout the viewing session. After viewing, participants completed a survey and were compensated. Figure 1 provides a

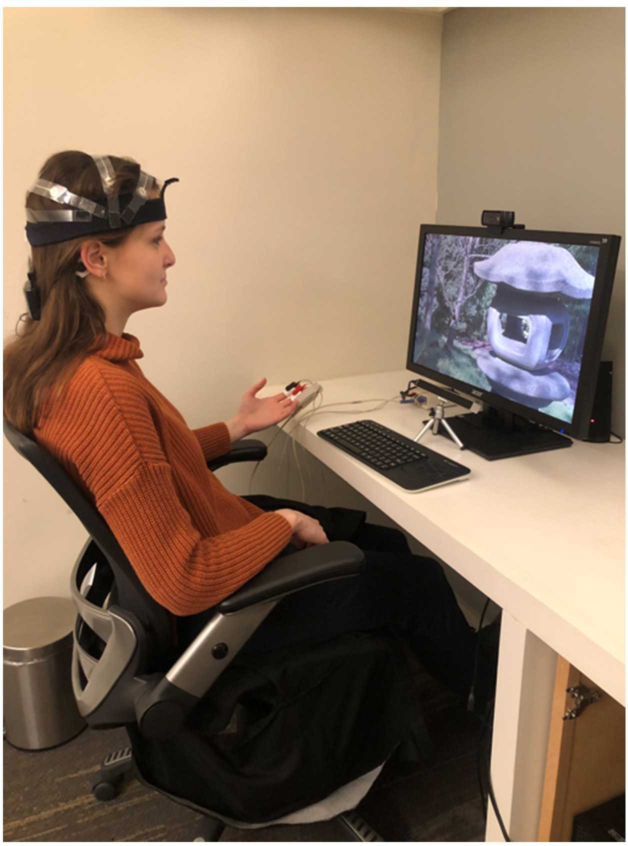

Figure 1. Lab experiment environment.

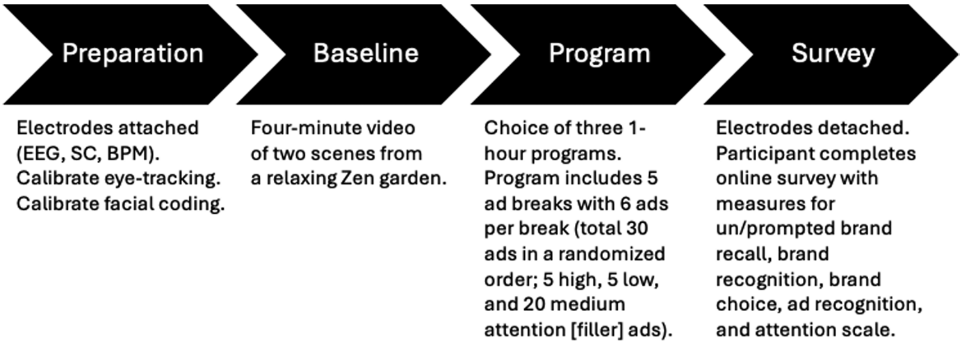

Figure 2. Lab test procedure.

photograph of the testing environment, and Figure 2 presents an overview of the experimental procedure.

The viewing session began with a four-minute meditative Zen Garden video, during which participants' resting baseline level was captured for all physiological measures. This baseline video was carefully chosen to observe resting baseline levels across all measures. Emotional arousal can take longer than three-minutes to subside to resting baseline level (Mattes and Cantor 1982), so the baseline video needed to be at least four-minutes long. Another potential baseline task might have been 'eyes closed,' but that state is associated with very high levels of EEG alpha waves (Mulholland 1973), which would invalidate change from baseline as a signature of attention using EEG. Evidence for the success of this baseline task is that average resting heart rate (RHR), measured during this task, was not significantly different from the average RHR reported in the most recent general public survey in the USA (95% confidence interval [73, 76] included $M = 73$, Aladin et al. 2014).

Participants then watched one-hour of television content, which included five breaks of six ads (i.e. 10 test ads plus 20 filler ads). Filler ads, also 30-seconds long, were from different categories to the test ads to avoid competitive interference effects (e.g. for the memory measures). Test and filler ads were inserted across the five breaks in an individually randomised order, to control for serial position and timing effects.

Participants were given a choice of programs (a drama, sitcom, or reality program) to equalize program liking, a proxy for mood (Petty et al. 1993), which could affect attention to the ads (Easterbrook 1959). All three programs put participants in a positive mood, as program liking (measured by a single 7-point item) was significantly higher than the scale mid-point for each program (M_{drama} = 6.5, M_{sitcom} = 6.1, $M_{reality}$ = 6.1). However, there were significant differences in program liking between the three programs (ANOVA $p = 0.02$; drama vs sitcom, $p = 0.01$; drama vs reality, $p = 0.03$; sitcom vs reality, $p = 0.8$). Importantly, these differences in mood did not result in significant differences in self-reported attention to the test ads between programs, collected with the post-viewing questionnaire. Self-reported ad attention averaged 62% during the drama, 59% during the sitcom, and 61% during the reality program (ANOVA $p = 0.66$). There were also no significant differences between programs in average ad attention responses according to the physiological measures (all p values $\geqq$ 0.08 [lowest p-value was for eyes-on-screen]).

Physiological measures

Physiological measures were collected via iMotions software, which synchronizes data from multiple sensors using the same annotated timeline, showing the beginning and end of each ad. Eye movements were tracked by EyeTech VT3 infra-red trackers. Heart rate and skin conductance were measured from three electrodes on the first three fingers of the nondominant hand. One electrode was a photoplethysmogram sensor that shone LED light on the skin to detect changes in light absorption, associated with blood pulsing. The other electrodes measured changes in the flow of a tiny electric current passing from one electrode to the other across the skin. Consenting participants had facial expressions recorded by a high-definition video camera mounted above the screen, with data computer analysed by the facial expression detection module of iMotions (AFFDEX) (McDuff and Kaliouby 2017). Participants wore an electrode cap with nine electrode sites to capture the EEG signal. Further measurement details are outlined in the online appendix.

Survey measures

Several intermediate advertising effectiveness measures were collected to confirm that attention was associated with better memory and potentially higher purchase likelihood, as well as the same self-reported attention scale used to select the ads. The measures, in order, were:

- Unprompted brand recall: Participants were asked to type all the brands they remembered seeing or hearing ads for during their session in separate text boxes. Text responses were reviewed and coded as Yes (1) or No (0) for each ad.
- Category prompted brand recall: Participants were asked “Do you recall the specific brands advertised for [insert category]?” for the target categories and to write the brand/s in an open-ended textbox. Text responses were reviewed and coded as Yes (1) or No (0) for each ad.
- Visual ad recognition: Participants were shown three images of key scenes from each ad that omit brand visuals and asked: “Do you remember seeing these commercials recently (including today’s session)?” Responses were coded as Yes (1) or No (0) for each ad.
- Self-reported attention: Participants were re-shown the same images and asked to indicate how much attention they had paid throughout the ad from 0% to 100%.
- Brand choice: Participants were asked to identify their preferred brand of two rival brands in each category (i.e., they had seen advertising for one of these brands and not the other).

Lift scores were calculated by comparing scores collected for the same brands with an unexposed control sample ($N = 145$). For the control sample, participants were invited to the lab to participate in a decoy task and completed the same survey even though they did not view the test ads.

Analysis

In a context where industry players are reporting 'black box' machine-learning measures of attention, we present model-free, causal evidence of the effects of experimentally manipulated levels of attention. We confirmed that the randomized order of presenting the test ads meant that there was no need to use order of presentation in a mixed-model regression. Hence, we report simple pairwise *t*-tests of individuals' difference between their average for the resting baseline video and average for each test ad (see Table 2).

Table 2. Means and paired *t*-test results versus resting baseline.

Ad/Mean	PSRA	SRA	BPM	EOS	Blink Dur.	FPS	IAD	Phasic SC	SCL	Smiling	EEG Alpha
Baseline			74	86	0.16	0.96	6.29	0.88	1.46	−1.01	4.03
			(11)	(19)	(0.03)	(0.49)	(0.37)	(0.21)	(0.52)	(1.46)	(0.08)
Low_1	55	39	75	89	0.16	1.31	6.11	0.90	1.49	−0.25	4.02
			(11)	(21)	(0.04)	(0.60)	(0.58)	(0.20)	(0.50)	(1.75)	(0.08)
			1.52	**2.13**	0.99	**8.90**	**4.55**	**2.00**	1.55	**9.44**	1.52
			NS	**0.02**	0.16	**<0.001**	**<0.001**	**0.02**	0.06	**<0.001**	0.07
Low_2	60	27	75	89	0.16	1.3	6.35	0.90	1.49	−0.36	4.03
			(11)	(20)	(0.04)	(0.64)	(0.57)	(0.21)	(0.50)	(1.65)	(0.09)
			2.91	**2.02**	0.68	**8.74**	1.53	**2.16**	1.63	**8.83**	0.03
			NS	**0.02**	NS	**<0.001**	NS	**0.02**	0.05	**<0.001**	NS
Low_3	62	44	73	89	0.15	1.44	5.97	0.90	1.48	−0.3	4.02
			(11)	(23)	(0.04)	(0.54)	(0.64)	(0.21)	(0.50)	(1.79)	(0.08)
			2.64	**2.14**	**2.00**	**12.27**	**7.44**	1.51	1.35	**7.91**	**2.71**
			<0.001	**0.02**	**0.02**	**<0.001**	**<0.001**	0.07	0.09	**<0.001**	**<0.001**
Low_4	59	27	74	89	0.16	1.25	6.05	0.91	1.5	−0.25	4.02
			(11)	(22)	(0.04)	(0.62)	(0.5)	(0.20)	(0.50)	(1.71)	(0.09)
			1.19	**1.75**	1.25	**7.77**	**6.77**	**2.53**	**2.21**	**10.32**	0.76
			NS	**0.04**	0.11	**<0.001**	**<0.001**	**0.01**	**0.01**	**<0.001**	0.22
Low_5	61	39	75	90	0.15	1.56	6.15	0.90	1.49	−0.26	4.02
			(11)	(22)	(0.04)	(0.66)	(0.57)	(0.21)	(0.51)	(1.80)	(0.08)
			2.10	**2.37**	**2.20**	**13.72**	**3.85**	**1.78**	1.56	**8.61**	1.37
			NS	**0.01**	**0.01**	**<0.001**	**<0.001**	**0.04**	0.06	**<0.001**	0.09
High_1	78	82	72	90	0.15	1.63	5.93	0.90	1.49	0.43	4.02
			(11)	(22)	(0.04)	(0.63)	(0.85)	(0.21)	(0.51)	(2.03)	(0.08)
			6.66	**2.80**	**1.89**	**15.97**	**6.49**	1.45	1.46	**14.05**	**2.49**
			<0.001	**<0.001**	**0.03**	**<0.001**	**<0.001**	0.07	0.07	**<0.001**	**0.01**
High_2	73	80	73	91	0.15	1.35	6.01	0.89	1.47	0.19	4.02
			(11)	(21)	(0.04)	(0.55)	(0.71)	(0.21)	(0.51)	(2.03)	(0.08)
			5.00	**3.42**	**2.37**	**10.22**	**6.06**	1.09	0.91	**11.35**	**2.90**
			<0.001	**<0.001**	**0.01**	**<0.001**	**<0.001**	0.14	0.18	**<0.001**	**<0.001**
High_3	78	87	73	90	0.16	1.44	5.97	0.89	1.47	−0.13	4.01
			(11)	(23)	(0.04)	(0.61)	(0.71)	(0.20)	(0.49)	(1.89)	(0.07)
			4.18	**2.42**	1.29	**11.17**	**6.86**	0.82	0.76	**9.42**	**4.88**
			<0.001	**0.01**	0.10	**<0.001**	**<0.001**	0.21	0.22	**<0.001**	**<0.001**
High_4	73	78	72	92	0.15	1.59	5.92	0.90	1.49	0.73	4.02
			(11)	(19)	(0.04)	(0.52)	(0.73)	(0.21)	(0.50)	(2.17)	(0.08)
			7.60	**4.07**	**2.82**	**15.74**	**7.42**	1.52	1.38	**14.60**	**3.17**
			<0.001	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.06	0.08	**<0.001**	**<0.001**
High_5	81	79	74	91	0.16	1.54	6.15	0.90	1.49	0.52	4.02
			(11)	(20)	(0.04)	(0.57)	(0.70)	(0.21)	(0.51)	(1.97)	(0.08)
			1.91	**3.04**	1.20	**14.85**	**2.87**	**1.92**	**1.76**	**14.33**	**3.43**
			0.03	**<0.001**	0.12	**<0.001**	**<0.001**	**0.03**	**0.04**	**<0.001**	**<0.001**
N			256	231	227	230	231	202	202	256	253

Note: **Bold = significant difference in the hypothesized direction** for ad mean versus the resting baseline mean (standard deviations in parentheses). Resting baseline = spent 4-min watching a Zen relaxation video prior to watching the program containing ads.

PSRA: pre-test self-reported attention; SRA: self-reported attention; BPM: beats per minute; IAD: individual attention dispersion; FPS: fixations per second; EOS: eyes on screen; SCL: skin conductance level; Phasic SC: phasic driver component of skin conductance; EEG: electroencephalogram.

Results

Manipulation checks confirmed that lab test participants rated the test ads similarly to the online survey participants ($r = 0.93$), with the low attention ads ($M = 58\%$, $SD = 3.2\%$) rated significantly lower for self-reported attention than the high attention ads ($M = 87\%$, $SD = 2.6\%$, $p < 0.001$).

Accuracy of attention measures

Given the challenges associated with a ground truth measure of attention, we triangulated results across measures. We use both classification performance (i.e. high attention vs low attention cf. resting baseline) (Table 3), and correlations with the EEG measure and self-reported attention scale (Table 4), to identify the potential validity of the physiological attention measures. Classification performance is evaluated using a confusion matrix. For example, for the SCL measure, the literature supported signature (i.e. significant positive change from resting baseline qualifies as attention) correctly classified four of five low attention ads (80%) as low attention, while it

Table 3. Percentage correct classification rates for 10 video ads.

	Classification vs resting baseline*			Classification vs 'average ad'**		
Measure	Correct Low Attention	Correct High Attention	Overall Accuracy	Correct Low Attention	Correct High Attention	Overall Accuracy
BPM	80	100	90	100	80	90
IAD	20	100	60	80	60	70
Blink Dur.	60	60	60	100	0	50
Smiling	0	100	50	100	80	90
FPS	0	100	50	80	60	70
EOS	0	100	50	100	20	60
SCL	80	20	50	80	0	40
Phasic SC	20	20	20	80	0	40
Average	33	75	54	90	38	64
EEG Alpha	80	100	90	80	80	80

Note: Table is ordered first by overall accuracy against resting baseline, then average ad baseline.
*Resting baseline = spent 4-min watching a Zen relaxation video prior to watching the program containing ads.
**'*Average Ad*' = average for measure over all 30 ads seen during the session.
BPM: beats per minute; IAD: individual attention dispersion; FPS: fixations per second; EOS: eyes on screen; SCL: skin conductance level; Phasic SC: phasic driver component of skin conductance, EEG: electroencephalogram.

Table 4. Correlations between attention measures.

	SRA	BPM	EOS	Blink	FPS	IAD	Ph. SC	SCL	Smile
SRA									
BPM	**−0.85**								
EOS	**0.74**	**−0.77**							
Blink Dur.	−0.49	**0.62**	**−0.59**						
FPS	**0.59**	**−0.55**	**0.60**	−0.43					
IAD	**−0.51**	**0.63**	−0.33	**0.57**	−0.36				
Phasic SC	**−0.85**	**0.75**	−0.49	**0.59**	−0.46	0.47			
SCL	**−0.63**	**0.60**	−0.35	0.27	−0.18	0.41	**0.64**		
Smiling	**0.73**	**−0.75**	**0.87**	−0.47	**0.67**	−0.48	−0.44	−0.13	
EEG Alpha	***−0.88***	***0.79***	***−0.65***	***0.54***	***−0.55***	*0.45*	***0.95***	***0.60***	***−0.61***

Note. Large correlations (absolute value > 0.5) shown in bold.
SRA: self-reported attention; BPM: beats per minute; EOS: eyes on screen; FPS: fixations per second; IAD: individual attention dispersion; Phasic SC: phasic driver component of skin conductance; SCL: skin conductance level; EEG: electroencephalogram.

correctly classified one of five high attention ads (20%) as high attention, meaning it accurately predicted 50% of the ads ([80% + 20%] ÷ 2 = 50%).

Only heart rate, specifically BPM, performed better than random chance at picking high attention ads from low attention ads, comparing each ad with the resting baseline (overall classification accuracy 90% vs 50% random chance, see Table 3). This supported H1. Heart rate was also highly correlated with self-reported attention (absolute value $r = 0.85$, see Table 4) and EEG ($r = 0.79$). The 'gold standard' but less scalable EEG measure had accuracy of 90%, which supported H9. Given all other measures performed as well as or worse than random chance, H2 through H8 were not supported.

The high positive correlation for phasic SC with EEG alpha ($r = 0.95$) (see Table 4), and its negative correlation with self-reported attention ($r = -0.85$), suggests that the *absence* of phasic SC could be an indicator of attention to ads. However, that result was contrary to H6. SCL also had low accuracy, and a negative correlation with self-reported attention, which contradicted H7.

The second-to-last row of Table 3 shows the average accuracy for distinguishing high from low attention across measures. For these theoretically derived comparisons with a resting baseline, the average classification performance was low and close to random chance at 54% accuracy. Averaging over 'apples and oranges' measures for attention accuracy may seem pointless but we deliberately used a toolbox of signatures that might capture different kinds of attention (e.g. negative vs positive emotion attention), and knowing that some measures are more responsive to different creative tactics (Bellman et al. 2019). And, as mentioned, a person's resting baseline is relevant for all measures. Consequently, it was surprising that average accuracy for high attention was very high (75%) while average accuracy for low attention was very low (33%). Average low attention accuracy was brought down by the eye tracking measures, particularly EOS and FPS, which misclassified *all* low attention ads as high attention. These results suggest people were not returning to their resting baseline state of visual attention when watching low attention ads.

Exploring an alternative baseline

Considering the poor low attention accuracy for multiple measures using resting baseline, responses during each ad were compared with attention during the 'average ad', based on all 30 ads seen. Using the 'average ad' baseline (see Table 5 for pairwise *t*-tests and Table 3 for classification performance), in addition to heart rate, smiling was revealed as another promising measure (accuracy 90%). Two eye tracking measures, IAD and FPS, also did significantly better than chance (accuracy 70%).

Benefits of attention

Table 6 reports the effects of ads with high versus low attention on survey measures of advertising effectiveness. Measuring lift from a control sample helps to control for differences in brand usage, as both were randomly recruited samples with equal or similar numbers of brand users/non-users. Using headroom lift, as opposed to absolute

Table 5. Means and paired *t*-test results versus the 'average ad'.

Ad/Mean	PSRA	SRA	BPM	EOS	Blink Dur.	FPS	IAD	Phasic SC	SCL	Smiling	EEG Alpha
'Average ad'			73	90	0.16	1.43	6.08	0.90	1.49	−0.06	4.02
			(11)	(17)	(0.03)	(0.46)	(0.51)	(0.2)	(0.49)	(1.65)	(0.08)
Low_1	55	39	75	89	0.16	1.31	6.11	0.90	1.49	−0.25	4.02
			(11)	(21)	(0.04)	(0.60)	(0.58)	(0.20)	(0.50)	(1.75)	(0.08)
			5.26	0.90	1.36	4.61	1.59	0.83	0.03	3.01	0.83
			NS	NS	NS	NS	NS	0.20	NS	NS	NS
Low_2	60	27	75	89	0.16	1.3	6.35	0.90	1.49	−0.36	4.03
			(11)	(20)	(0.04)	(0.64)	(0.57)	(0.21)	(0.50)	(1.65)	(0.09)
			6.87	1.24	3.74	4.19	8.33	1.42	0.31	5.27	2.82
			NS	NS	NS	NS	NS	0.08	0.38	NS	NS
Low_3	62	44	73	89	0.15	1.44	5.97	0.90	1.48	−0.3	4.02
			(11)	(23)	(0.04)	(0.54)	(0.64)	(0.21)	(0.50)	(1.79)	(0.08)
			0.81	0.63	0.46	0.48	**2.04**	0.43	0.42	4.49	**1.80**
			0.21	NS	NS	NS	**0.02**	NS	NS	NS	**0.04**
Low_4	59	27	74	89	0.16	1.25	6.05	0.91	1.5	−0.25	4.02
			(11)	(22)	(0.04)	(0.62)	(0.5)	(0.20)	(0.50)	(1.71)	(0.09)
			4.35	1.61	0.75	7.31	0.20	**2.07**	**1.66**	3.85	1.74
			NS	NS	NS	NS	0.42	**0.02**	**0.05**	NS	NS
Low_5	61	39	75	90	0.15	1.56	6.15	0.90	1.49	−0.26	4.02
			(11)	(22)	(0.04)	(0.66)	(0.57)	(0.21)	(0.51)	(1.80)	(0.08)
			5.24	0.17	0.07	**3.79**	3.26	0.41	0.20	3.77	1.05
			NS	NS	0.47	**<0.001**	NS	0.34	0.42	NS	NS
High_1	78	82	72	90	0.15	1.63	5.93	0.90	1.49	0.43	4.02
			(11)	(22)	(0.04)	(0.63)	(0.85)	(0.21)	(0.51)	(2.03)	(0.08)
			5.73	0.27	0.13	**7.89**	**3.54**	0.50	0.06	**7.63**	1.38
			<0.001	0.4	NS	**<0.001**	**<0.001**	NS	NS	**<0.001**	0.08
High_2	73	80	73	91	0.15	1.35	6.01	0.89	1.47	0.19	4.02
			(11)	(21)	(0.04)	(0.55)	(0.71)	(0.21)	(0.51)	(2.03)	(0.08)
			4.10	0.97	1.40	3.93	1.61	1.63	1.62	**3.83**	**1.92**
			<0.001	0.17	0.08	NS	0.05	NS	NS	**<0.001**	**0.03**
High_3	78	87	73	90	0.16	1.44	5.97	0.89	1.47	−0.13	4.01
			(11)	(23)	(0.04)	(0.61)	(0.71)	(0.20)	(0.49)	(1.89)	(0.07)
			2.63	0.11	0.61	0.003	**2.60**	2.37	2.01	1.47	**4.92**
			<0.001	NS	NS	NS	**<0.001**	NS	NS	NS	**<0.001**
High_4	73	78	72	92	0.15	1.59	5.92	0.90	1.49	0.73	4.02
			(11)	(19)	(0.04)	(0.52)	(0.73)	(0.21)	(0.50)	(2.17)	(0.08)
			8.32	**2.40**	1.41	**6.78**	**4.93**	0.45	0.38	**10.38**	**2.72**
			<0.001	**0.01**	0.08	**<0.001**	**<0.001**	NS	NS	**<0.001**	**<0.001**
High_5	81	79	74	91	0.16	1.54	6.15	0.90	1.49	0.52	4.02
			(11)	(20)	(0.04)	(0.57)	(0.70)	(0.21)	(0.51)	(1.97)	(0.08)
			0.15	0.93	0.88	**4.31**	3.36	0.41	0.36	**8.57**	**2.95**
			NS	0.18	NS	**<0.001**	NS	0.34	0.36	**<0.001**	**<0.001**
N			256	231	227	230	231	202	202	256	253

Note: **Bold = significant difference in the hypothesized direction** for ad mean versus the 'average ad' mean (standard deviations in parentheses). '*Average ad*' = average for measure over all 30 ads seen during the session.

PSRA = pre-test self-reported attention, SRA = self-reported attention, BPM = beats per minute, IAD = individual attention dispersion, FPS = fixations per second, EOS = eyes on screen, SCL = skin conductance level, Phasic SC = phasic driver component of skin conductance, EEG = electroencephalogram.

Table 6. Brand lift versus unexposed control sample.

Headroom Lift	Low Attention Ads	High Attention Ads
Unprompted brand recall	10	25
Category prompted brand recall	17	38
Visual ad recognition	59	93
Brand choice	0	27

Note. Headroom lift = (experiment minus control)/(1 minus control).

or relative lift, further standardizes measurement across bigger and smaller brands. Lift, versus an unexposed control sample, was significantly different from zero for every measure except for low attention brand choice. Lift was also consistently higher for high attention ads (cf. low attention ads). These results demonstrate that attention matters for advertising, although it likely has diminishing returns.

Discussion

This research tested and compared scalable physiological measures for their ability to identify low and high attention ads, compared to a 'gold standard' non-scalable measure, EEG, and a self-reported attention measure.

Heart rate, measured as BPM, proved the best measure of attention to accurately classify ads as high versus low attention using its hypothesized signature comparison with the individual's resting baseline. It also performed well using a different baseline, the individual's 'average ad'. It was as accurate as EEG Alpha. Thus, heart rate is supported as a measure that can accurately classify high and low attention to video advertising, supporting other research which demonstrates its utility in marketing measurement (Gangadharbatla, Bradley, and Wise 2013; Simmonds et al. 2020b).

All other measures (EOS, blink duration, FPS, IAD, phasic SC, SCL and smiling) proved inaccurate because they consistently misclassified low attention ads according to our hypotheses. For these measures, it appears even low attention advertising attracts attention compared to resting baseline, particularly visual attention. However, this conflicted with consumers' ratings, as they perceived differences between low versus high attention ads. These results prompted exploration into a potential ad-specific baseline, which has no basis in theory but may benefit applied research. Using this ad-specific baseline, BPM was again the strongest scalable measure, but smiling was also promising, followed by IAD and FPS.

These findings highlight potential issues with using only one measure of attention as a currency for advertising.

Theoretical implications

Building on earlier systematic comparisons of measures (e.g. Venkatraman et al. 2015) this research contributes to the literature on attention to video advertising by comparing different measures of attention, all of which were expected to distinguish between high and low attention to video advertising, as rated by consumers. Consumers appear to associate high attention with high/er levels of conscious processing, in line with dual processing theories (Chaiken 1980; Petty and Cacioppo 1986). For this reason, the measures that aligned better with how consumers rated attention were those that more directly measured conscious processing, such as BPM (Lang 1994), as opposed to automatic monitoring (Koch and Tsuchiya 2007; Krauzlis et al. 2023; MacInnis and Jaworski 1989; Oberauer 2019). We show that when some attention measures are indicating high attention, there may be only moderate, low, or no conscious ad processing occurring, as measured by EEG and BPM. Visual attention measures are particularly responsive to automatic scanning processes that can involve no conscious ad processing or emotional engagement (Anderson, Bothell, and Douglass 2004; Hawkins et al. 1997; Koch and Tsuchiya 2007; MacInnis and Jaworski 1989). Because ads include many features that automatically capture attention (Wooley et al. 2022), visual attention measures can mislead about the level of conscious attention these ads receive (Rosenholtz 2024). However, visual attention measures could be good choices for measuring automatic attention unrelated to memory and other aspects of conscious processing (Wedel and Pieters 2000, 2017).

The lack of support for most of our hypotheses raises questions about whether these new results falsify previous studies, or whether our study was a poor test of those hypotheses. The best answer is that most of the studies we relied on did not explicitly test a change from resting baseline hypothesis. For example, Heath, Nairn, and Bottomley (2009) measured change from baseline with FPS, but their baseline was attention to the program, rather than a relaxation video. In their study, FPS during ads could be higher *or lower* than baseline, like our 'average ad' baseline. However, Heath, Nairn, and Bottomley (2009) compared ads with ads, rather than ads with baseline to draw their conclusions. Similarly, the literature on blink duration (Morris and Miller 1996; Tijerina et al. 1999) tested changes over time but never change from baseline. Other measures controlled for individual differences without using change from baseline, and compared ads with ads, rather than ads with baseline (e.g. smiling: McDuff et al. 2015; IAD: Teixeira, Wedel, and Pieters 2010). We suggest the failures of our hypothesis tests do not mean that the literature supporting them is no longer supported. Instead, our results suggest these measures fail to discriminate between conscious and unconscious attention.

The results for phasic SC and SCL are more troubling because the skin conductance literature recommends and tests change from baseline hypotheses (Potter and Bolls 2012). Also, increases in skin conductance are indicators of sympathetic nervous system arousal and therefore emotional and perhaps conscious responding to ads (Koruth et al. 2015). We note that our sample of ads were likely associated with smaller arousal responses than the loud noises used in the validation of phasic SC (Benedek and Kaernbach 2010). These results should not be used to argue that ads generally are not very arousing, as prior studies have reported high arousal during ads (e.g. Bolls, Muehling, and Yoon 2003). Future research testing phasic SC and SCL signatures of attention should use ads likely to generate a strong arousal response (e.g. fear ads: Lee and Lang 2009).

Our results support suggestions that asking respondents is a direct, reliable way to measure attention (Krauzlis et al. 2023), although self-reported 'attention' also appears to measure perceived ad processing effort. Prior literature criticizes self-reported attention measures (e.g. Atalay, Bodur, and Rasolofoarison 2012; Kolar et al. 2021), mainly because self-reports do not capture unconscious attention. We argue that self-reported attention is valuable precisely because it does reflect conscious processing. For example, our smiling measure was unable to distinguish high from low attention ads compared with resting baseline, because our test ads consistently used a positive emotional tone. But our participants were able to distinguish between these ads, perhaps because they knew when they were consciously smiling as opposed to unconsciously smiling (Ekman 1991). Our survey question was correlated with measures of both attention and ad processing (EEG and BPM). However, dual processing theories (Chaiken 1980; Petty and Cacioppo 1986) explain why advertising can also be effective in conditions of unconscious low attention (e.g. when there is no evidence of ad recognition, Santoso et al. 2022).

Industry implications

This research contributes to practice by identifying measures that can most reliably detect both high and low attention to video advertising. It has lessons for advertisers and agencies, media platforms and research providers. Our results suggest that distinguishing high attention from low attention, as consumers perceive this difference, requires measures of conscious ad processing as well as attention, and heart rate appears to do both. However, we acknowledge that a toolbox of measures can have a role given that advertisers are also interested in the effectiveness of low attention and unconscious ad exposures.

In the current discussions about an attention currency, this research highlights that there are alternatives to eye tracking that offer accurate measurement, which will be critical if media are priced on attention. It supports using heart rate as a validated measure of high versus low visual attention, which can also measure attention to sound (Simmonds et al. 2020b). Heart rate can potentially be measured at scale (e.g. as a currency measure), across ads and media including radio, using webcams or wearable technologies (Nahler et al. 2018).

A key difficulty when measuring heart rate in-field is obtaining individuals' resting baselines. Fortunately, a resting state consistent with that measured during our baseline video is a common state to which people return, which means individuals' baselines could be measured, for example, during relaxing nature documentaries (Nelson, Meyvis, and Galak 2009), or with known resting-state (low attention) ads, like those included in this research.

If some advertisers merely wish to distinguish between more versus less attention, our results show that several measures can be used with a data-driven 'average ad' baseline. Consenting viewers in large panels could each have an individual 'average ad' baseline calibrated from their responses to thousands of ads. Small differences from this baseline would be statistically significant, and as an indication of how useful this baseline might be, shifts in the hypothesized direction (i.e. ignoring significance) were 80% accurate for blink duration and EOS (cf. 50% and 60% accurate respectively, as reported in Table 3 at $p < 0.05$), which are two easily scalable visual attention measures.

Limitations and future research

Despite being the most comprehensive comparison of physiological attention measures applied to video advertising, to date, this study has limitations. For example, we omitted other attention measures, such as fEMG, pupil dilation, and average fixation duration, which future research could pursue. More validation testing is also encouraged, with different samples of consumers and ads, across different media platforms. Future research could also investigate using more levels of attention, going beyond our dichotomy of high versus low attention, to include deeper measures of ad processing, such as elaboration (Greenwald and Leavitt 1984).

We needed a lab test to include EEG as a gold-standard comparison measure, but this raises legitimate concerns about forced exposure, and heightened attention levels (e.g. no multitasking). But even in this forced exposure condition, ads were seen with low attention, as rated by participants. For visual attention, EOS averaged around 90%, not 100%, over the session, which indicates people were mostly but not exclusively looking at the screen. Importantly, heart rate identified low attention even when people were mostly looking at the screen (i.e. when eye tracking measures might register 'full attention'). It is possible that in real-world distracted viewing situations that simpler measures, like EOS, could be good enough measures of attention. We also encourage validation studies triangulating multiple measures in more realistic viewing settings (e.g. in-home or co-viewing scenarios).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by Google. However, both the Ehrenberg-Bass Institute and MediaScience also receive funding from Google's competitors (e.g. television networks and social media). This balance of interests means that the authors' views are not influenced by any single company or industry group.

Data availability statement

Physiological data that support the main findings of this study, which were collected by MediaScience, are available via Open Science Foundation at https://osf.io/6fsm5/.

References

Aladin, A.I., S.P. Whelton, M.H. Al-Mallah, M.J. Blaha, S.J. Keteyian, S.P. Juraschek, J. Rubin, C.A. Brawner, and E.D. Michos. 2014. Relation of resting heart rate to risk for all-cause mortality by gender after considering exercise capacity (the henry ford exercise testing project). *The American Journal of Cardiology* 114, no. 11: 1701–6.

Alba, J.W, and J.W. Hutchinson. 1987. Dimensions of consumer expertise. *Journal of Consumer Research* 13, no. 4: 411–54.

Anderson, J.R., D. Bothell, and S. Douglass. 2004. Eye movements do not reflect retrieval processes: Limits of the eye-mind hypothesis. *Psychological Science* 15, no. 4: 225–31.

Atalay, A.S., H.O. Bodur, and D. Rasolofoarison. 2012. Shining in the center: Central gaze Cascade effect on product choice. *Journal of Consumer Research* 39, no. 4: 848–66.

Barrett, L.F. 2022. Facial expressions do not reveal emotions. Scientific American, https://www.scientificamerican.com/article/darwin-was-wrong-your-facial-expressions-do-not-reveal-your-emotions/

Barrett, L.F., R. Adolphs, S. Marsella, A.M. Martinez, and S.D. Pollak. 2019. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest: A Journal of the American Psychological Society* 20, no. 1: 1–68.

Bazzani, A., S. Ravaioli, L. Trieste, U. Faraguna, and G. Turchetti. 2020. Is EEG suitable for marketing research? A systematic review. *Frontiers in Neuroscience* 14: 594566.

Bellman, S., M. Nenycz-Thiel, R. Kennedy, L. Larguinat, B. McColl, and D. Varan. 2017. What makes a television commercial sell? Using biometrics to identify successful ads. *Journal of Advertising Research* 57, no. 1: 53–66.

Bellman, S., M. Nenycz-Thiel, R. Kennedy, N. Hartnett, and D. Varan. 2019. Best measures of attention to creative tactics in TV advertising: When do attention-getting devices capture or reduce attention? *Journal of Advertising Research* 59, no. 3: 295–311.

Benedek, M., and C. Kaernbach. 2010. A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods* 190, no. 1: 80–91.

Boksem, M.A.S, and A. Smidts. 2015. Brain responses to movie trailers predict individual preferences for movies and their population-wide commercial success. *Journal of Marketing Research* 52, no. 4: 482–92.

Bolls, P.D., A. Lang, and R.F. Potter. 2001. The effects of message valence and listener arousal on attention, memory, and facial muscular responses to radio advertisements. *Communication Research* 28, no. 5: 627–51.

Bolls, P.D., D.D. Muehling, and K. Yoon. 2003. The effects of television commercial pacing on viewers' attention and memory. *Journal of Marketing Communications* 9, no. 1: 17–28.

Chaiken, S. 1980. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology* 39, no. 5: 752–66.

Chang, Y., and E. Thorson. 2023. Media multitasking, counterarguing, and brand attitude: Testing the mediation effects of advertising attention and cognitive load. *Computers in Human Behavior* 139: 107544.

Coulson, S., and M. Kutas. 2001. Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters* 316, no. 2: 71–4.

Couwenberg, L.E., M.A.S. Boksem, R.C. Dietvorst, L. Worm, W.J.M.I. Verbeke, and A. Smidts. 2017. Neural responses to functional and experiential ad appeals: Explaining ad effectiveness. *International Journal of Research in Marketing* 34, no. 2: 355–66.

Cowan, N. 1995. *Attention and memory: An integrated framework*. New York: Oxford University Press.

Dehaene, S. 2014. *Consciousness and the brain: Deciphering how the brain codes our thoughts*. New York: Viking.

Dentsu. 2021. Unlocking the new currency of attention. https://www.dentsu.com/uk/en/reports/unlocking_the_new_currency_of_attention_fa35bf9

Easterbrook, J. 1959. The effect of emotion on cue utilization and the organization of behavior. *Psychological Review* 66, no. 3: 183–201.

Ekman, P. 1991. *Telling lies: Clues to deceit in the marketplace, politics and marriage*. US: W.W. Norton & Company.

Ekman, P., and W.V. Friesen. 1978. *The facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Gangadharbatla, H., S. Bradley, and W. Wise. 2013. Psychophysiological responses to background brand placements in video games. *Journal of Advertising* 42, no. 2-3: 251–63.

Gevins, A.S., G.M. Zeitlin, J.C. Doyle, C.D. Yingling, R.E. Schaffer, E. Callaway, and C.L. Yeager. 1979. Electroencephalogram correlates of higher cortical functions. *Science (New York, N.Y.)* 203, no. 4381: 665–8.

Gidlöf, K., A. Wallin, R. Dewhurst, and K. Holmqvist. 2013. Using eye tracking to trace a cognitive process: Gaze behaviour during decision making in a natural environment. *Journal of Eye Movement Research* 6, no. 1: 1–14.

Gordon-Hecker, T., A. Pittarello, S. Shalvi, and M. Roskes. 2020. Buy-one-get-one-free deals attract more attention than percentage deals. *Journal of Business Research* 111: 128–34.

Graham, F.K, and R.K. Clifton. 1966. Heart-rate change as a component of the orienting response. *Psychological Bulletin* 65, no. 5: 305–20.

Graham, F.K. 1997. Afterword: Pre-attentive processing and passive and active attention. In *Attention and orienting: Sensory and motivational processes*, eds. Lang P.J., Simons F.P. and Balaban M., 417–52. Erlbaum Hillsdale, NJ: Erlbaum.

Greenwald, A.G, and C. Leavitt. 1984. Audience involvement in advertising: Four levels. *Journal of Consumer Research* 11, no. 1: 581–92.

Hartnett, N., R. Kennedy, B. Sharp, and L. Greenacre. 2016. Creative that sells: How advertising execution affects sales. *Journal of Advertising* 45, no. 1: 102–12.

Hawkins, R.P., S. Pingree, L. Bruce, and J. Tapper. 1997. Strategy and style in attention to television. *Journal of Broadcasting & Electronic Media* 41, no. 2: 245–64.
Heath, R. 2000. Low involvement processing—A new model of brands and advertising. *International Journal of Advertising* 19, no. 3: 287–98.
Heath, R. 2009. Emotional engagement: How television builds big brands at low attention. *Journal of Advertising Research* 49, no. 1: 62–73.
Heath, R., A.C. Nairn, and P.A. Bottomley. 2009. How effective is creativity? Emotive content in TV advertising does not increase attention. *Journal of Advertising Research* 49, no. 4: 450–63.
Heath, R., D. Brandt, and A. Nairn. 2006. Brand relationships: Strengthened by emotion, weakened by attention. *Journal of Advertising Research* 46, no. 4: 410–9.
James, W. 1890. *Principles of psychology*. Cambridge, MA: Harvard University Press.
Janiszewski, C. 1993. Preattentive mere exposure effects. *Journal of Consumer Research* 20, no. 3: 376–92.
Kakaria, S., E. Bigné, V. Catrambone, and G. Valenza. 2023. Heart rate variability in marketing research: A systematic review and methodological perspectives. *Psychology & Marketing* 40, no. 1: 190–208.
Kim, J., S. Kang, and K.H. Lee. 2021. Evolution of digital marketing communication: Bibliometric analysis and network visualization from key articles. *Journal of Business Research* 130: 552–63.
Koch, C., and N. Tsuchiya. 2007. Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences* 11, no. 1: 16–22.
Kolar, T., Z. Batagelj, I. Omeragić, and M. Husić-Mehmedović. 2021. How moment-to-moment EEG measures enhance ad effectiveness evaluation: Peak emotions during branding moments as key indicators. *Journal of Advertising Research* 61, no. 4: 365–81.
Koruth, K.J., A. Lang, R.F. Potter, and R.L. Bailey. 2015. A comparative analysis of dynamic and static indicators of parasympathetic and sympathetic nervous system activation during TV viewing. *Communication Methods and Measures* 9, no. 1–2: 78–100.
Krauzlis, R.J., L. Wang, G. Yu, and L.N. Katz. 2023. What is attention? *Wiley Interdisciplinary Reviews. Cognitive Science* 14, no. 1: E 1570.
Krugman, H.E. 1965. The impact of television advertising: Learning without involvement. *Public Opinion Quarterly* 29, no. 3: 349–56.
Lacey, J.I. 1967. Somatic response patterning and stress: Some revisions of activation theory. In *Psychological stress: Issues in research*, ed. M.I. Appley and R. Trumbull, 14–42. New York, NY: Applleton-Centry-Crofts.
Lajante, M.M.P., O. Droulers, and D. Amarantini. 2017. How reliable are "state-of-the-art" facial EMG processing methods? Guidelines for improving the assessment of emotional valence in advertising. *Journal of Advertising Research* 57, no. 1: 28–37.
Lang, A. 1990. Involuntary attention and physiological arousal evoked by structural features and emotional content in TV commercials. *Communication Research* 17, no. 3: 275–99.
Lang, A. 1994. What can the heart tell us about thinking?. In *Measuring psychological responses to media messages*, ed. A. Lang, 99–111. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
Lang, A., J. Borse, K. Wise, and P. David. 2002. Captured by the world wide web: Orienting to structural and content features of computer-presented information. *Communication Research* 29, no. 3: 215–45.
Lee, S., and A. Lang. 2009. Discrete emotion and motivation: Relative activation in the appetitive and aversive motivational systems as a function of anger, sadness, fear, and joy during televised information campaigns. *Media Psychology* 12, no. 2: 148–70.
Loschky, L.C., A.M. Larson, J.P. Magliano, and T.J. Smith. 2015. What would jaws do? The tyranny of film and the relationship between gaze and higher-level narrative film comprehension. *PloS One* 10, no. 11: E0142474.
Macinnis, D.J., C. Moorman, and B.J. Jaworski. 1991. Enhancing and measuring consumers' motivation, opportunity, and ability to process brand information from ads. *Journal of Marketing* 55, no. 4: 32–53.
Mack, A., and I. Rock. 1998. *Inattentional blindness MIT press/Bradford books series in cognitive psychology*. Cambridge, MA: MIT Press.

MacInnis, D.J, and B.J. Jaworski. 1989. Information processing from advertisements: Toward an integrative framework. *Journal of Marketing* 53, no. 4: 1–23.
Mattes, J., and J. Cantor. 1982. Enhancing responses to television advertisements via the transfer of residual arousal from prior programming. *Journal of Broadcasting* 26, no. 2: 553–66.
McDuff, D., and R.E. Kaliouby. 2017. Applications of automated facial coding in media measurement. *IEEE Transactions on Affective Computing* 8, no. 2: 148–60.
McDuff, D., R.E. Kaliouby, J.F. Cohn, and R.W. Picard. 2015. Predicting ad liking and purchase intent: Large-scale analysis of facial responses to ads. *IEEE Transactions on Affective Computing* 6, no. 3: 223–35.
Media Rating Council. 2019. *MRC cross-media audience measurement standards (phase. 1 video)*, 1–63. United States: Media Rating Council.
Morris, T.L, and J.C. Miller. 1996. Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology* 42, no. 3: 343–60.
Mulholland, T. 1973. Objective EEG methods for studying covert shifts of visual attention. In *The psychophysiology of thinking*, eds. FJ Mcguigan and RA Schoonover New York: Academic Press.
Nahler, C., B. Feldhofer, M. Ruether, G. Holweg, and N. Druml. 2018. Exploring the usage of time-of-flight cameras for contact and remote photoplethysmography. Paper presented at the 21st Euromicro Conference on Digital System Design (DSD), August 2018.
Nelson, L.D., T. Meyvis, and J. Galak. 2009. Enhancing the television-viewing experience through commercial interruptions. *Journal of Consumer Research* 36, no. 2: 160–72.
Oberauer, K. 2019. Working memory and attention – A conceptual analysis and review. *Journal of Cognition* 2, no. 1: 36.
Oelofse, A., S. Malherbe, I.S. Pretorius, and M. Du Toit. 2010. Preliminary evaluation of infrared spectroscopy for the differentiation of Brettanomyces bruxellensis strains isolated from red wines. *International Journal of Food Microbiology* 143, no. 3: 136–42.
Öhman, A. 1997. As fast as the blink of an eye: Evolutionary preparedness for preattentive processing of threat. In *Attention and orienting: Sensory and motivational processes*, eds. P.J. Lang, R.F. Simons and M. Balaban, 165–84. Hillsdale, NJ: Erlbaum.
Orquin, J.L., M.P. Bagger, and S.M. Loose. 2013. Learning affects top down and bottom up modulation of eye movements in decision making. *Judgment and Decision Making* 8, no. 6: 700–16.
Petty, R.E, and J.T. Cacioppo. 1986. *Communication and persuasion*. New York: Springer-Verlag.
Petty, R.E., D.W. Schumann, S.A. Richman, and A.J. Strathman. 1993. Positive mood and persuasion: Different roles for affect under high- and low-elaboration conditions. *Journal of Personality and Social Psychology* 64, no. 1: 5–20.
Petty, R.E., G.L. Wells, and T.C. Brock. 1976. Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology* 34, no. 5: 874–84.
Petty, R.E., J.T. Cacioppo, and D. Schumann. 1983. Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research* 10, no. 2: 135–46.
Potter, R.F., and P.D. Bolls. 2012. *Psychophysiological measurement and meaning*. New York, NY: Routledge.
Ramsøy, T.Z. 2019. Building a foundation for neuromarketing and consumer neuroscience research: How researchers can apply academic rigor to the neuroscientific study of advertising effects. *Journal of Advertising Research* 59, no. 3: 281–94.
Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124, no. 3: 372–422.
Rees, G., C. Russell, C.D. Frith, and J. Driver. 1999. Inattentional blindness versus inattentional amnesia for fixated but ignored words. *Science (New York, N.Y.)* 286, no. 5449: 2504–7.
Reeves, B., E. Thorson, M.L. Rothschild, D. Mcdonald, J. Hirsch, and R. Goldstein. 1985. Attention to television: Intrastimulus effects of movement and scene changes on alpha variation over time. *The International Journal of Neuroscience* 27, no. 3–4: 241–55.

Regan, M.A., C. Hallett, and C.P. Gordon. 2011. Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident; Analysis and Prevention* 43, no. 5: 1771–81.

Rosenholtz, R. 2024. Visual attention in crisis. *The Behavioral and Brain Sciences:* 1–32. https://doi.org/10.1017/S0140525X24000323

Russell, J.A, and L.F. Barrett. 1999. Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology* 76, no. 5: 805–19.

Santoso, I., M.J. Wright, G. Trinh, and M. Avis. 2022. Mind the attention gap: How does digital advertising impact choice under low attention? *European Journal of Marketing* 56, no. 2: 442–66.

Segijn, C.M., H. Voorveld, L. Vandeberg, and E.G. Smit. 2017. The battle of the screens: Unraveling attention allocation and memory effects when multiscreening. *Human Communication Research* 43, no. 2: 295–314.

Shiffrin, R.M, and W. Schneider. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84, no. 2: 127–90.

Simmonds, L., S. Bellman, R. Kennedy, M. Nenycz-Thiel, and S. Bogomolova. 2020a. Moderating effects of prior brand usage on visual attention to video advertising and recall: An eye-tracking investigation. *Journal of Business Research* 111: 241–8.

Simmonds, L., S. Bogomolova, R. Kennedy, M. Nenycz-Thiel, and S. Bellman. 2020b. A dual-process model of how incorporating audio-visual sensory cues in video advertising promotes active attention. *Psychology & Marketing* 37, no. 8: 1057–67.

Simons, R.F., B.H. Detenber, B.N. Cuthbert, D.D. Schwartz, and J.E. Reiss. 2003. Attention to television: Alpha power and its relationship to image motion and emotional content. *Media Psychology* 5, no. 3: 283–301.

Sokolov, E.N, and J.T. Cacioppo. 1997. Orienting and defense reflexes: Vector coding the cardiac response. In *Attention and orienting: Sensory and motivational processes*, ed. P.J. Lang, R.F. Simons, and M. Balaban Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Stern, J.A., D. Boyer, and D. Schroeder. 1994. Blink rate: A possible measure of fatigue. *Human Factors* 36, no. 2: 285–97.

Sung, B., N.J. Wilson, J.H. Yun, and E.J. Lee. 2020. What can neuroscience offer marketing research? *Asia Pacific Journal of Marketing and Logistics* 32, no. 5: 1089–111. https://doi.org/10.1108/APJML-04-2019-0227

Teixeira, T., M. Wedel, and R. Pieters. 2010. Moment-to-moment optimal branding in TV commercials: Preventing avoidance by pulsing. *Marketing Science* 29, no. 5: 783–804.

The Attention Council. 2021. *From attention to action*. New York: The Attention Countil.

Thorson, E. 1994. Using eyes on screen as a measure of attention to television. In *Measuring psychological responses to media messages*, ed. A. Lang, 65–84. Hillsdale, NJ: Lawrence Erlbaum Associates.

Tijerina, L., W.W. Wierwille, M.J. Goodman, S. Johnston, D. Stoltzfus, and M. Gleckler. 1999. *A preliminary assessment of algorithms for drowsy and inattentive driver detection on the road*. Washington, DC: National Highway and Safety Administration, Department of Transport.

Varan, D., A. Lang, P. Barwise, R. Weber, and S. Bellman. 2015. How reliable are neuromarketers' measures of advertising effectiveness? *Journal of Advertising Research* 55, no. 2: 176–91.

Varan, D., M. Nenycz-Thiel, R. Kennedy, and S. Bellman. 2020. The effects of commercial length on advertising impact: What short advertisements can and cannot deliver. *Journal of Advertising Research* 60, no. 1: 54–70.

Venkatraman, V., A. Dimoka, P.A. Pavlou, K. Vo, W. Hampton, B. Bollinger, H.E. Hershfield, M. Ishihara, and R.S. Winer. 2015. Predicting advertising success beyond traditional measures: New insights from neurophysiological methods and market response modeling. *Journal of Marketing Research* 52, no. 4: 436–52.

Wainer, H. 1991. Adjusting for differential base rates: Lord's paradox again. *Psychological Bulletin* 109, no. 1: 147–51.

Weber, R., J.M. Mangus, and R. Huskey. 2015. Brain imaging in communication research: A practical guide to understanding and evaluating fMRI studies. *Communication Methods and Measures* 9, no. 1-2: 5–29.

Wedel, M., and R. Pieters. 2000. Eye fixations on advertisements and memory for brands: A model and findings. *Marketing Science* 19, no. 4: 297–312.

Wedel, M., and R. Pieters. 2017. A review of eye-tracking research in marketing. In *Review of marketing research* (Vol. 4), ed. N.K. Malhotra. Leeds: Emerald Group Publishing Limited.

Westermann, J.-F., R. Schäfer, M. Nordmann, P. Richter, T. Müller, and M. Franz. 2024. Measuring facial mimicry: Affdex vs. EMG. *PloS One* 19, no. 1: E0290569.

Wooley, B., S. Bellman, N. Hartnett, A. Rask, and D. Varan. 2022. Influence of dynamic content on visual attention during video advertisements. *European Journal of Marketing* 56, no. 13: 137–66.

Zajonc, R.B. 1968. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology* 9, no. 2, Pt.2: 1–27.