Predicting advertising effectiveness by facial expressions in response to amusing persuasive stimuli

Lewinski, P.; Fransen, M.L.; Tan, E.S.H.

Published in:
Journal of Neuroscience, Psychology, and Economics

DOI:
10.1037/npe0000012

Journal of Neuroscience, Psychology, and Economics. 10.1037/npe0000012

Predicting advertising effectiveness by facial expressions in response to amusing persuasive stimuli.

Lewinski, P., Fransen, M. L., Tan, E.S.H.

Date submitted: 15 March 2013
Date resubmitted: 31 May 2013
Date accepted: 12 August 2013

Predicting Advertising Effectiveness by Facial Expressions in Response to Amusing Persuasive Stimuli

Peter Lewinski, Marieke L. Fransen and Ed S.H. Tan.

The Amsterdam School of Communication Research ASCoR, Department of Communication Science, University of Amsterdam

An extended abstract of this paper was published in *2013 NeuroPsychoEconomics Conference Proceedings* (pp. 45-45). Bonn, Germany: Association for NeuroPsychoEconomics.

We would like to thank Tim den Uyl, Paul Ivan and Marten den Uyl, Hanneke Koster and Mariska Snijdewind for the valuable feedback received on the earlier versions of this manuscript.

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement 290255.

Peter Lewinski works - as a Marie Curie Research Fellow - for Vicarious Perception Technologies B.V. – an artificial intelligence company that develops FaceReader software for Noldus Information Technologies B.V. He is also a PhD Candidate in ASCoR under supervision of prof. dr. Ed S.H. Tan and dr. Marieke L. Fransen

Correspondence concerning this article should be addressed to Peter Lewinski, The Amsterdam School of Communication Research ASCoR, Department of Communication Science, University of Amsterdam, East Indies House (OIH) Kloveniersburgwal 48, Amsterdam 1012 CX. E-mail: p.lewinski@uva.nl

Abstract

We present a psychophysiological study of facial expressions of happiness (FEH) produced by advertisements using the FaceReader system (Noldus, 2013) - for automatic analysis of facial expressions of basic emotions (FEBE) (Ekman, 1972). FaceReader scores were associated with self-reports of the advertisement's effectiveness. Building on work describing the role of emotions in marketing research, we examined the relationship between the patterns of the FEBE and the perceived amusement of the advertisements, attitude toward the advertisement (AAD) and attitude toward the brand (AB). Differences were observed between FEH scores in response to highly, medium, and low amusing video advertisements (AVAs). Positive correlations were found between FEH and AAD and FEH and AB in high and medium but not in low AVAs. As hypothesized, other basic emotions (sadness, anger, surprise, fear and disgust) did not predict advertisement amusement or advertisements' effectiveness. FaceReader enabled a detailed analysis of more than 120,000 frames of video-recordings contributing to an identification of global patterns of facial reactions to amusing persuasive stimuli. For amusing commercials, context-specific FEH features were found to be the major indicators of advertisement effectiveness. The study used video-recordings of participants in their natural environments obtained through a crowd-sourcing platform. The naturalistic design of the study strengthened its ecological validity and demonstrated the robustness of the software algorithms even under austere conditions. Our findings provide first evidence for the applicability of FaceReader methodology in the basic consumer science research.

Keywords: facial expressions of emotion, AAD, AB, FaceReader, amusement

Predicting Advertising Effectiveness by Facial Expression in Response to Amusing Persuasive Stimuli

The ultimate goal of advertising and marketing strategies is to deliver persuasive communication convincing another party to change their opinion or attitude (Meyers-Levy & Malaviya, 1999). The senders of persuasive message – e.g. advertising and marketing companies – engage in active listening (e.g. customer panels) in order to constantly gauge customers' opinions along with their current and future attitudes. To gain better insight into customers' behavior, it is possible to simply ask what they think and what they think they feel. However asking requires quite an effort and brings along undesired effects such as self-awareness (Pryor, Gibbons, Wicklund, Fazio & Hood, 1977) and social-desirability (Arnold & Feldman, 1981) in the questioned participants. Persuasive agents would want to know more about "sincere" acts and observe and quantify behavior that is not easy moldable by the person under investigation. Advertising and marketing companies are looking for ultimate tools that can assess and predict the behavior, and that the customers cannot "hide" their answer.

Catering for such unfulfilled wishes, the subfield of neuromarketing has been flourishing in recent years. Researchers used (a) brain imaging (Langleben et al. 2009); (b) EEG (e.g. Ohme, Reykowska, Wiener & Choromanska, 2009; Cook, Warren, Pajot, Schairer & Leuchter, 2011); (c) electrodermal response registration (for a review see Lajante, Droulers, Dondaine, & Amarantini, 2012); (d) eye tracking (Wedel & Pieters, 2000; Pieters & Wedel, 2004; Ramsøy, Friis-Olivarius, Jacobsen, Jensen & Skov, 2012); (e) heart rate registration (Micu & Plummer, 2010); and (f) facial analysis (Teixeira, Wedel & Pieters, 2012) to assess effectiveness of their advertising and marketing campaigns. The present study employed automated facial analysis applied to a sample of 270 participants' facial reactions to persuasive amusing stimuli. The like sample would virtually defy human coding for uncovering global patterns of facial behavior, if

only for reasons of time and money. FaceReader (Noldus, 2013) – the facial analysis software we used – analyzed, on frame-by-frame basis more than 8,000 seconds of video recordings, i.e. around 120,000 frames on six basic emotions scales; an amount of material that would likely take months for two or three independent human coders to analyze.

Measuring Emotions in Advertising

Marketers believe that emotions are an important aspect of consumer behavior in a persuasive context. Wiles and Cornwell (1990) some time ago already reviewed the tools that were used to measure emotions in advertising research. Poels and Dewitte (2006) provided an update distinguishing the following measures (a): self-report: verbal, visual, moment-to-moment; and (b) autonomic: heart rate, skin conductance, facial expressions. Researchers can use these tools to assess traditional dependent variables that capture the effects of advertisements (a) attitude toward the advertisement (AAD), (b) attitude toward the brand (AB), and (c) purchase intention (PI). In their review, Poels and Dewitte (2006) concluded first that emotions can predict advertisements' effects and second that autonomic measures of emotions have higher predictive power than self-reports. However, in practice autonomic measures are used less commonly due to the high-cost sophisticated research set-up they require.

Self-reports

Emotion self-reports capture subjective feelings of respondents. Verbal self-report involves either answering open-ended questions, rating experienced emotions on a scale or both. Two major approaches - “dimensional” and “basic emotions” – respectively represent emotions as (a) positioned on three independent bipolar dimensions (Olney, Holbrook & Batra, 1991) or (b) a blend of basic emotions from a limited set (e.g. Zeitlin & Westwood, 1986). The three dimensions - pleasure, arousal and dominance – jointly capture more information on immediate reactions to advertisements, than each of the so-called basic emotions as e.g. happiness, surprise

or sadness (Havlena & Holbrook, 1986). Visual self-reports ask to rate emotional states by choosing a cartoon character representing emotion felt. AdSAM® (Morris, Woo, Geason & Kim, 2002) and PrEmo (Desemet, 2002) - two prime examples - are taken faster and are less boring to respondents than their verbal counterparts (Poels & Dewitte, 2006). Moment-to-moment reports ("feelings monitors") capture valence - positive vs. negative - of experienced emotions immediately and in-real time (Baumgarter, Sujan & Padgett, 1997). Self-reports have been popular in advertisement research because they are cheap, quick, user-friendly, valid and efficient. However, their validity has major shortcomings, too. First, they fail to capture low-order emotions, i.e. ones resulting from low-complexity automatic processes such as pleasure and arousal. In addition, they elicit socially desirable answers and tend to increase "cognitive bias" (Poels & Dewitte, 2006). It seems that autonomic measures provide solutions for these limitations.

Autonomic Measures

Autonomic measures capture the bodily reactions that are often beyond the person's conscious control. It is the autonomic nervous system that is mainly in control of those physiological reactions (Winkielman, Berntson & Cacioppo, 2001). The Facial Action Coding System (Ekman & Friesen, 1978; Ekman, Friesen & Hager, 2002) is a widely used scientific tool to describe visible movement of facial muscles. Unfortunately, it did not prove sensitive enough to measure emotional reactions to advertisements (Derbaix, 1995). Facial electromyography (EMG) fares better in that, it correlates with self-report measures. Zygomatic and corrugator muscle activity - smiling and frowning respectively – have been shown to capture the valence of emotions (Lang, Greenwald, Bradley & Hamm, 1993). Facial EMG is an invasive tool though due to placement of electrodes on the participants' face, which decreases ecological validity. Skin conductance (SC) is a measure capturing less noise but it requires extensive training and

knowledge of advanced statistics. Changes in SC indicate autonomic nervous system activation: the higher SC levels the higher is physiological arousal (Ravaja, 2004). Nevertheless, the valance of the emotional stimuli cannot be determined because either positive or negative stimuli would evoke the same increased SC activation values (Hopkins & Fletcher, 1994). Heart rate and heart beat patterns can indicate arousal, valence and attention in real-time and continuous manner. However, because it captures a variety of phenomena it is recommended for use in tandem with other measures e.g. SC to secure ecological validity (Hopkins & Fletcher, 1994).

Advertisement Effectiveness

The core elements of advertisement effectiveness are the attitude toward the advertisement (AAD) and attitude toward the brand (AB), each measuring slightly different concepts with AAD likely contributing to the formation of AB (Mitchell & Olson, 1982). AAD captures people's liking, enjoyment and the valence of their feelings toward the commercial whereas AB captures those toward the brand advertised in the commercial (Chattopadhyay & Basu, 1990; Phillips, 2000). Scholars often include also purchase intention and actual buying behavior in the measurements of advertisement effectiveness but those two concepts seem hard to capture by self-reports measures and need more sophisticated methods to study them. The current study measured AAD and AB and related them to particular patterns of facial expressions. Thus we wanted to establish the first link between people's facial behavior and their attitudes towards advertisements and brands and hence advertisement effectiveness in the amusing video ads.

FaceReader

In order to overcome some of the limitations of the most commonly used tools in the advertising research the emotional reactions to advertisements can be measured by autonomic responses namely facial expressions. The feasibility of the approach is augmented when an automated tool is used. This study employed a commercially available, advanced and unobtrusive

tool provided by Noldus – FaceReader, version 5.0 to capture and analyze facial expressions of emotions. FaceReader has been used in a variety of contexts fit for experimental research in consumer behavior. Some recent studies featuring FaceReader include automated facial analysis of expressions elicited by orange juices (Danner, Sidorkina, Joechl, & Duerrschmid, 2013), (dis)liked food (de Wijk, Kooijman, Verhoeven, Holthuyzen, & Graaf, 2012; He, Boesveldt, de Graaf, & de Wijk, 2012) and effects of facial emotional feedback on readiness to use computer-based assessment (Terzis, Moridis, & Economides, 2012; 2013). In their 2012 study Terzis, Moridis and Economides measured behavioral intentions using FaceReader. To our knowledge there is no research testing FaceReader as a tool to measure advertisement effectiveness operationalized as self-reported AAD and AB.

Facial Expressions

Facial expressions reflect affective states defined for example in EMFACS-7 (Friesen & Ekman, 1983) and therefore possibly predict associated behavior and attitude change. Facial expressions of emotions are semi-universal sequences of facial muscle contractions linked with the emotional state of the person. The neurocultural theory of emotion, which is advocated by Ekman (e.g. Ekman, 1972; Ekman & Cordano, 2011), defines facial expressions of emotion as discrete, innate and culturally independent.

According to other researchers, there is a two-way link between facial expressions and emotion regulation (Cole, 1986; Gross & Thompson, 2007; Izard, 1990). This is why in research on facial expressions it is difficult to establish causal relationships between facial nonverbal behavior and interpretations assigned to them - the emotions. Emotions do cause facial expressions (“I feel happy so I smile”), but facial expressions also cause emotions (“I smile and it makes me happy”). Any causal relationship between smiling and perception of the advertisement

has not been established in the advertisement context. Smiling or laughing may indicate liking for the advert and therefore the advertisement's higher effectiveness.

Amusing Advertisements

This study assessed whether people's facial reactions to video advertisements differ depending on how amusing the persuasive stimuli are. The more amusing the stimuli the more laughing and smiling they should elicit. In order to test the generality of the relation between FEH and amusement two checks were built into the design of the study. First, the amusement potential of the stimulus ads was varied enabling one to check whether the predicted relationship holds for different amusement levels. Second, it is necessary to check for other emotions than happiness being predictive of the amusement of the stimuli. There is a recent evidence that the FEH is the most reliable and robust emotion of all FEBE, see a meta-analysis by Reisenzein, Studtmann and Horstmann (2013). Nonetheless it could be possible that a stimulus is not amusing but rather positively surprising or that a low – amusing stimulus elicits also facial expressions of negative emotions (e.g. sadness, anger, disgust) due to its lacking enjoyableness. Thus, we test:

H_1: Highly amusing video advertisements elicit more frequent and more intense facial expressions of happiness than medium and low amusing ones;

H_2: There is no difference between highly, medium and low amusing video advertisements in the frequency and intensity of facial expressions of all other basic emotions but happiness.

We hypothesize that the aggregated patterns of facial expressions of emotions help in predicting people's attitudes. People experiencing positive emotions seeing an ad are more likely to have positive attitudes toward the advertisement (AAD) and the brand (AB). Moreover, if H_1

and H_2 are true then the FEH should indicate the effectiveness of the more amusing video advertisements but not of the low amusing ones, thus we also test:

$H_{3:}$ There is a positive correlation between frequency and intensity of facial expressions of happiness and self-reported measures of ad effectiveness (a) attitude toward the ad and (b) attitude toward the brand in the highly and medium but not in the low amusing video advertisements;

$H_{4:}$ There is no correlation between frequency and intensity of facial expressions and all other basic emotions but happiness and self-reported measures of ad effectiveness (a) attitude toward the ad and (b) attitude toward the brand in any of the highly, medium or low amusing video advertisements.

Video advertisements were pre-tested as to how amusing they were before facial reactions and questionnaires data were collected for the test of the four hypotheses.

The Experimental Research

Pretesting

The selection of amusing video advertisements (AVAs) started with three experts (specialized advertising researchers from our lab) electing 16 ads they judged to differ as to how amusing the ads were. In a next step equal length ads (M= 30 sec., SD = 2 sec.) were retained. The ads presented well- and less- known U.S. brands, included both services and products, and targeted both females and males. Test participants were recruited through Amazon Mechanical Turk (MTurk; Buhrmester, Kwang, & Gosling, 2011) who were U.S. residents and native English speakers. An average number of 30 females and 30 males per advertisement rated how amusing the ads were using the following items (α = 0.89) each scored one to five: funny (1= *not-funny*, 5

= *funny)*; perceived intensity of smiling (1 = *not at all*, 5 = *to a great extent);* perceived intensity of laughing (1 = *not at all* and 5 = *to a great extent)*. Participants were asked if they had seen the advertisement or brand previously. In the end, six ads were selected, two for each condition of high, medium, and low AVAs that received highest, medium and lowest scores respectively from the initial pool of 16 video advertisements. Amusement scores and additional characteristics for these six selected advertisements can be found in Table 1.

[Table 1 about here]

Table 1

Amusement Scores and Characteristics of Selected Video Advertisements

Advertisement	*M*	*SD*	*n*	Brand	Product	Utility	Target-group	Amusement
Free Doritos	3.85	1.21	60	Doritos	Chips	Good	General	High
E-Trade Baby Girlfriend	3.78	1.06	68	E-Trade	E-trade	Service	General	High
GEICO - Dancing Kitten	2.93	1.26	63	GEICO	Auto insurance	Service	General	Medium
Dr. Pepper Ten	2.80	1.17	62	Dr. Pepper	Soft drink	Good	Males	Medium
Teleflora Adriana Lima	2.25	1.12	64	Teleflora	Flowers delivery	Service	Males	Low
Wonderstruck Taylor Swift	2.03	1.05	60	Taylor Swift	Perfumes	Good	Females	Low

There were no significant differences between the two ads from each of the conditions, the genders nor between the participants who saw already the advertisement or brand before.

Study Design

Participants. A fresh sample consisting of 90 participants (51 Men, 39 Women, average age = 27.14 years, *SD*=9.05) - were recruited through MTurk. Through credit card and IP check it was secured that - participants could participate only once, were U.S. residents, native English speakers and older than 18 years. In order to participate, a person had to agree to an informed consent. In addition, participants had to have a standard computer with a webcam and the Flash

plug-in correctly installed. The effective hourly rate was on average 5.08 $ and participants took on average six minutes and 50 seconds (*SD* =77 sec.) to finish the experiment.

Stimuli. The stimuli - the six video advertisements that were chosen in the pretesting - aimed to be funny and amusing, i.e. evoking smiles and laughs. There were two ads for each group of high, medium and low AVAs with global average respective amusement scores M_h = 3.81, SD_h = 1.12; M_m = 2.86, SD_m = 1.21; M_l = 2.14, SD_l = 1.08.

Design and procedure. The entire data collection was done through MTurk. If the person accepted the task, they were redirected to an external server where the experiment was executed. Each participant was exposed to the three different high, medium, and low AVAs in a random order. Participants were recorded through their own webcams while watching the video advertisements. They were aware of and agreed to be video-recorded. Before each recording, they were instructed to position their head in the center of the webcam's focus, reminded to maintain that posture and keep their hands on the keyboard during the entire recording. Participants were further informed that they had to watch three video advertisements and answer three alike questionnaires about the ad seen after each trial. They were not given any additional instructions, and were debriefed in the end.

Measures

Objective (FaceReader; den Uyl & van Kuilenburg, 2005) - and subjective (self-reports; Chattopadhyay & Basu, 1990; Phillips, 2000) - measurements were used to capture reactions to persuasive amusing stimuli. Control measures included gender and product-category involvement (four items, α = 0.88) influences.

FaceReader – FEBE. This system uses a 3-layer neural network that automatically recognizes and analyzes facial expressions of emotions in humans (den Uyl & van Kuilenburg, 2005). It classifies the facial expressions from pictures and videos into the following categories of

basic emotions (a) happiness; (b) sadness; (c) anger; (d) surprise; (e) fear; (f) disgust (cf. Ekman, 1972). FaceReader, at a first stage, finds a face using the Active Template Method. Then it creates a virtual, super-imposed 3D Active Appearance Model of the face featuring almost 500 unique landmarks. In a third stage, scores for the intensity and probability of facial expressions for basic emotions are computed (van Kuilenburg, Wiering, & den Uyl, 2005). The neural network of the system has been trained using a high-quality set of approximately 10,000 images that were manually annotated by the human coders. The average rates of performance reported are 89% (den Uyl & van Kuilenburg, 2005; van Kuilenburg, Wiering, den Uyl, 2005) and 87% (Terzis, Moridis, & Economides, 2013).

Self-reports of advertisements effectiveness. Advertisement effectiveness was defined as an (a) attitude toward the ad (AAD) and (b) attitude toward the brand (AB). Attitudes were measured by three 5-point Likert scale items. AAD (α = 0.96) had three items from Phillips (2000): - I think the commercial that I just watched is *bad* (=1) - *good (=5); unlikable* (=1) - *likeable* (=5); *not enjoyable* (=1) - *enjoyable* (= 5). For AB (α = 0.91) a scale composed of three items was adopted from Chattopadhyay and Basu (1990): *bad* (=1) - g*ood* (=5)*; unlikable* (=1) – *likeable* (=5)*; negative* (=1) – *positive* (=5).

Results

FaceReader– output analysis. The system assigned to each frame of the video recording of facial reactions an estimation of the intensity of facial expression of six basic emotions from 0 to 1. The global mean average score of the top 10% peak values of facial expressions of emotions to perform all the calculations was used. To compute the value for one of the emotions for each participant their facial expression scores for that emotion were ordered from lowest to highest. The relative frequency of each of the scores was computed and also the cumulative relative frequency for each subsequent score, resulting in a percentile distribution. When for each

participant all values below the 90th percentile are removed, the analysis will be based on a top 10 % criterion, when all scores under the 80th percentile are removed on a top 20% criterion etc. For each participant the average score of the top 10% peak values for all facial expressions of emotions were calculated. Finally, the global mean for each facial expression of emotion using the average scores defined above was computed. We chose this approach in order to analyse the most prominent facial expressions and take into account the frequency of their occurrence during exposure.

Additionally, calculations of 20% and 100% top peak values were made by removing all the values below the 80th percentile or using all the values, respectively. They were used to check whether the cut-off criterion influenced the results, which it appeared they did not. Since FaceReader scores were not normally distributed non-parametric Friedman and Wilcoxon tests were used for comparison of medians and Spearman coefficients in the computation of correlations. See Figure 1 for an example of graphic output of the FaceReader.

[Figure 1 about here]

Figure 1. FaceReader's Graphic Output – Participant 9

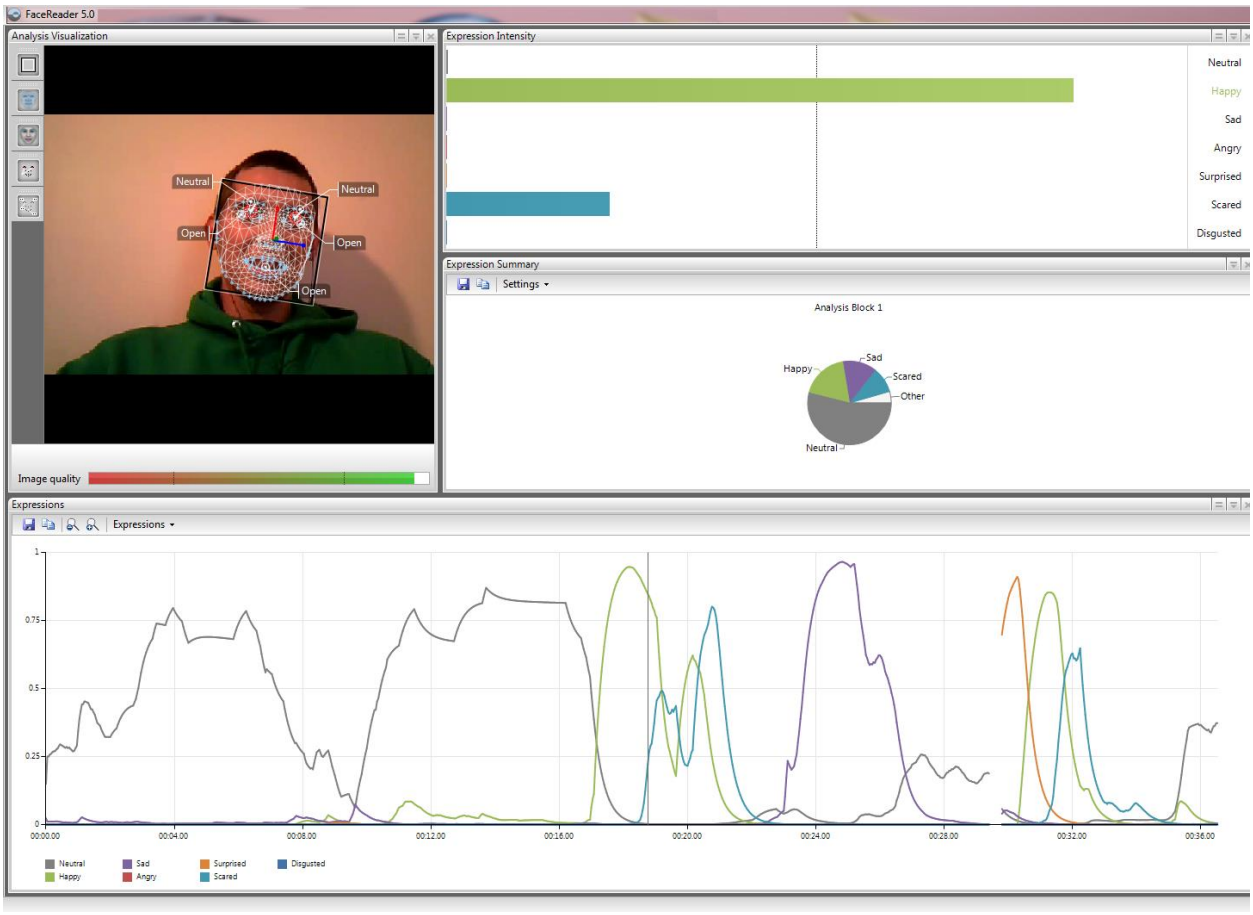


Figure 1. Plots of facial reactions to highly amusing video advertisements. Different colors represent basic emotions. Published with written permission.

Amusement of advertisements. In order to test H_1 the facial expressions of happiness were analyzed; to test H_2 expressions of the other basic emotions were analyzed. Friedman tests were run to determine differences in the global mean average score of top 10% peak values of facial expressions of happiness (FEH-10%) and the scores differed, $\chi 2(2) = 31.40$, $p < .001$ (see Table 2). Pairwise comparisons with Bonferroni correction for multiple comparisons showed that scores were higher in the high, (*Mdn* = 0.66) than in the medium, (*Mdn* = 0.39), $z = 3.05$, $p = .002$ and low, (*Mdn* =.10), $z = 5.78$, $p < .001$ and in the medium than in the low, $z = 4.12$, $p < .001$

AVAs conditions. The global mean average score of top 10% peak values of facial expressions of all other basic emotions were not different in any of the high, medium and low AVAs conditions, as summarized in Table 2.

[Table 2 about here]

Table 2

Global Average Scores for All Other Emotions in Different Amusement Conditions

	Median					
Emotion	H	M	L	$\chi 2(2)$	*p*	*n*
Happiness	**.66**	**.39**	**.10**	***31.40***	***<.001***	*90*
Sadness	.24	.23	.23	.07	.97	90
Anger	.06	.08	.13	2.22	.33	90
Surprise	.00	.00	.01	*	*	90
Fear	.00	.00	.00	*	*	90
Disgust	.00	.00	.00	*	*	90

Note., H/M/L = high, medium and low AVAs conditions; Emotion = facial expressions of basic emotion (Ekman, 1972) as defined in FaceReader (den Uyl & van Kuilenburg, 2005); * Friedman test cannot be performed due to insufficient variance of scores.

Advertisement effectiveness. In order to test H_3, first the facial expressions of happiness and then to test H_4, the facial expressions of all other basic emotions in relation to the advertisement effectiveness were analyzed. See Table 3 for average self-reported advertisement effectiveness scores in the high, medium and low AVAs.

[Table 3 about here]

Table 3

Average Scores of Self-reported Advertising Effectiveness

	Highly-amusing		Medium-amusing		Low-amusing		
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*n*
AAD	4.03	1.19	3.38	1.32	3.29	1.14	90
AB	4.00	.99	3.80	.98	3.50	1.12	90

Note., AAD = attitude toward ad; AB = attitude toward brand

Spearman's rank-order correlations were calculated to assess the relationship between facial expressions of happiness and advertisement effectiveness. Table 4 summarizes the results. Figure 2 offers a graphic representation of the correlations. In the high and medium AVAs there was a positive correlation between FEH-10% and AAD, respectively $r_s(90) = .61$, $p < .001$; $r_s(90) = .28$, $p < .001$ and between FEH-10% and AB, respectively $r_s(90) = .49$, $p < .001$; $r_s(90) = .26$, $p < .001$. No correlations obtained in the low AVAs condition.

[Table 4 about here]

Table 4

Correlation between FEH-10% and AAD & AB

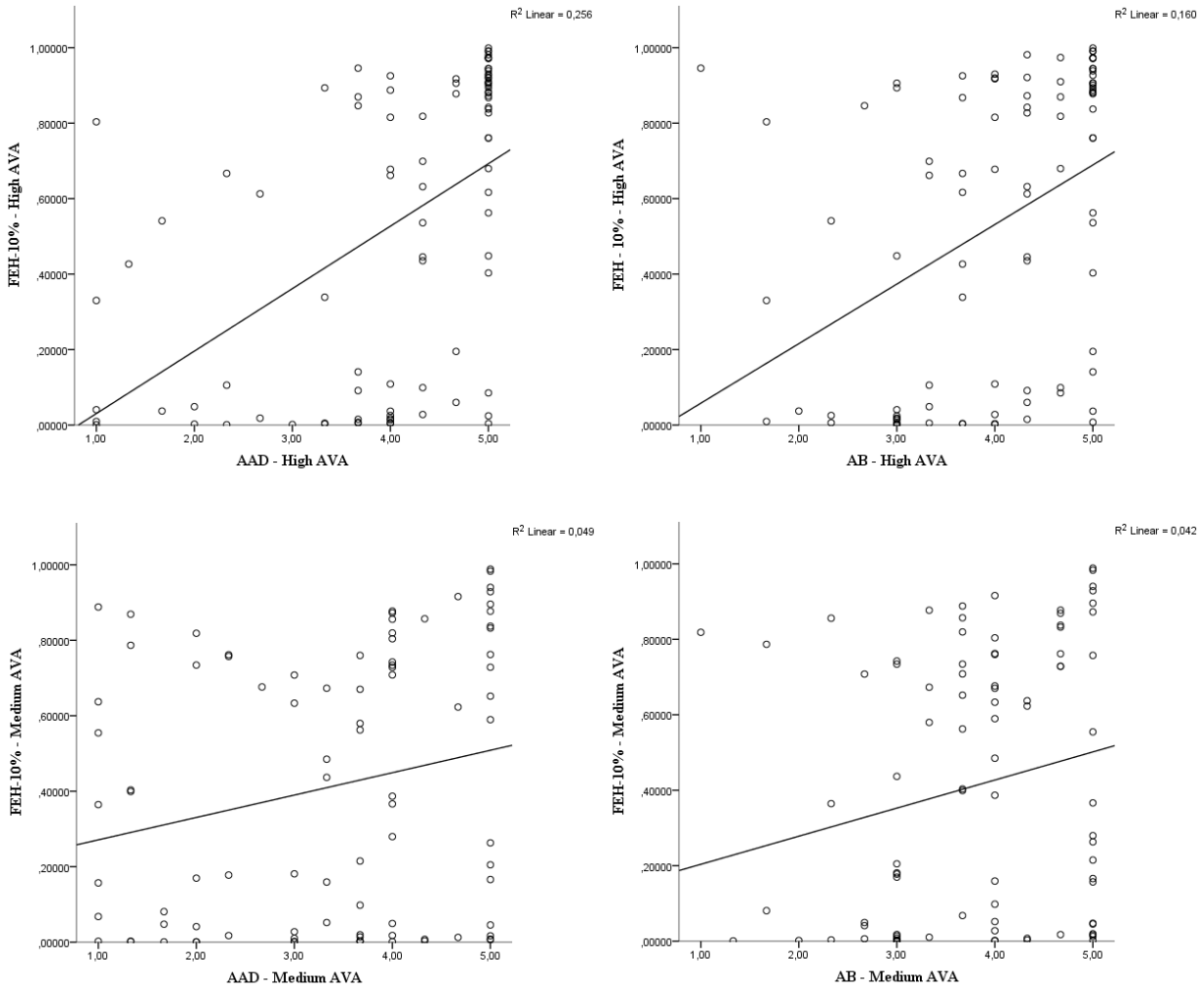
	AAD		AB		
AVAs	r_s	*p**	r_s	*p**	*n*
High	**.61**	**<.001**	**.49**	**<.001**	90
Medium	**.28**	**<.001**	**.26**	**<.001**	90
Low	.11	.32	.04	.69	90

Note. AAD = attitude toward advertisement; AB = attitude toward brand; AVAs = amusing video advertisements

*two-tailed.

[Figure 2 about here]

Figure 2. Facial Expressions of Happiness and Advertising Effectiveness



Scatterplots of FEH-10% and AAD in high and medium AVA conditions (left panels) and FEH-10% and AB in high and medium AVA conditions (right panels)

Spearman's rank-order correlations were used to assess the relationship between facial expressions of all other basic emotions and the advertisement effectiveness. No significant correlations were found between these variables in any of the high, medium and low AVAs conditions. See Table 5 for the summary of the results.

[Table 5 about here]

Table 5

Correlation between Facial Expressions of All Other Basic Emotions and AAD & AB

	Highly-amusing				Medium-amusing				Low-amusing				
	AAD		AB		AAD		AB		AAD		AB		
Emotion	r_s	*p*	r_s	*p*	r_s	*p*	r_s	*p*	r_s	*p*	r_s	*p*	*n*
Sadness	-.06	.56	.16	.88	.12	.25	-.17	.12	-.12	0.26	-.11	.32	90
Anger	-.02	.83	-.02	.86	-.13	.22	-.19	.08	-.02	.88	-.05	.65	90
Surprise	-.08	.45	-.06	.58	-.18	.27	.01	.96	.02	.82	-.05	.66	90
Fear	.11	.31	-.12	.26	-.10	.33	-.20	.06	.04	.71	-.05	.65	90
Disgust	.13	.24	-.05	.63	.08	.44	-.04	.71	.06	.56	.17	.10	90

Note, AAD = attitude toward ad and AB = attitude toward brand; Emotion = facial expressions of basic emotion (Ekman, 1972) as defined in FaceReader 5.0 (den Uyl & van Kuilenburg, 2005).

Additional analysis. Additional analysis were performed of the global mean average score on the basis of top 20% and 100% peak values of facial expressions to identify differences in amusement of the advertisements and in correlations with advertisement effectiveness. See Table 6 and Table 7 for the summary of the results.

The pattern of the results is almost the same as the one obtained analyzing the top 10 % peak values. There was one exception: using top 100% facial expression values the happiness score does not correlate significantly with AAD in the medium – AVA condition.

[Table 6 about here]

[Table 7 about here]

Table 6

Average Scores for All Emotions in Top Peak Values – 20% & 100% Approaches

	Median -20%					Median -100%				
Emotion	H	M	L	$\chi 2(2)$	*p*	H	M	L	$\chi 2(2)$	*p*
Happiness	.49	.24	.06	37.42	< .001*	.12	.09	.01	27.82	< .001**
Sadness	.18	.16	.16	.09	.96	.05	.04	.05	.36	.84
Anger	.04	.05	.08	2.52	.28	.01	.01	.02	1.36	.51
Surprise	.00	.00	.01	***	***	.00	.00	.00	***	***
Fear	.00	.00	.00	***	***	.00	.00	.00	***	***
Disgust	.00	.00	.00	***	***	.00	.00	.00	***	***

Note, 20% / 100% = the global mean average score of 20% / 100% top peak values of facial expressions of respective emotion; H/M/L = high , medium and low AVAs conditions; Emotion = facial expressions of basic emotion (Ekman, 1972) as defined in FaceReader 5.0 (den Uyl & van Kuilenburg, 2005); *We performed Pairwise comparisons (SPSS) with a Bonferroni correction for multiple comparisons and the scores were higher in the high, (*Mdn* = 0.49) than in the medium (*Mdn* = 0.24) , *z* = 3.36, *p* < .001 and low (*Mdn* =.06), *z* = 5.88, *p* < .001 and in the medium than in the low, *z* = 4.23, *p* < .001 AVAs condition; ** We performed Pairwise comparisons (SPSS) with a Bonferroni correction for multiple comparisons and the scores were higher in the high, (*Mdn* = 0.12) than in the medium, (*Mdn* = 0.09) , *z* = 2.10, *p* < .001 and low , (*Mdn* =.01), *z* = 5.35, *p* < .001 and in the medium , than in the low, *z* = 4.01, *p* < .001 AVAs condition; ***not possible to run Friedman test due to too little variance of the scores.

Table 7

Correlations for All Emotions in Top Peak Values – 20% & 100% Approaches

	20%						100%					
	AAD			AB			AAD			AB		
Emotion	H	M	L	H	M	L	H	M	L	H	M	L
Happiness	**.58**	**.25**	.12	**.47**	**.27**	.05	**.52**	.20	.12	**.40**	**.26**	.06
Sadness	-.09	.12	-.12	-.04	-.20	-.01	-.11	.10	-.13	-.06	-.21	-.07
Anger	-.02	-.13	.00	-.02	-.18	-.04	.04	-.13	.01	-.03	.-17	-.02
Surprise	-.08	-.12	.02	-.06	.01	-.05	-.09	-.10	.01	-.07	.02	-.07
Fear	.12	-.08	.01	-.11	-.19	-.05	.12	-.13	.01	-.10	-.08	-.07
Disgust	.13	.09	.05	-.05	.11	.17	.10	.08	.05	-.06	.00	.18

Note, 20%/100% = 20%/100% of the global mean average score of the top peak values of facial expressions of corresponding emotion; AAD = attitude toward ad; AB = attitude toward brand; Emotion = facial expressions of basic emotion (Ekman, 1972) as defined in FaceReader 5.0 (den Uyl & van Kuilenburg, 2005); H/M/L = high, medium, and low AVAs conditions.

In addition, across all conditions and regardless of top peak value criterion there was neither any difference between genders nor between participants reporting higher versus lower involvement in the product-category. There was no correlation between the FEH in high AVAs with AAD an AB in medium and low AVAs, neither between the FEH in medium AVAs with AAD and AB in low AVAs, demonstrating that people did not show FEH indicating random liking. A check on effects of order of presentation of advertisements did not yield significant effects either.

Limitations

The current study has some limitations that deserve address in further research. First, facial expressions may have been biased by various factors. For example, people who scored higher on facial expressions of happiness may have been in a generally better mood when entering the experiment and throughout. Control pre-measures were not used in order to prevent priming the participants before the recordings, while post-measures would have been confounded with experimental effects. Instead, we opted to recruit a sufficient sample of reactions and use repeated measures so as to minimize the influence of mood. As another example, people with different emotion regulation strategies (Gross, 2003) may have exaggerated or in contrast down-regulated their facial expressions in the presence of the persuasive stimuli depending on what they deemed appropriate (e.g. because of social desirability).

Moreover, characteristics of the preselected stimuli such as target-group, product vs. service or well– vs. less– known brands could potentially provoke different facial reactions. In addition, happy facial reactions could be due to presence of people in the selected stimuli – causing participants to simply mimic facial expressions of actors in the advertisements. We did not control for the above-mentioned variables because we were interested in overall affective reactions (i.e. amusement) the stimuli provoked and not the particular characteristics of those ads.

A limitation of another kind is the scope of materials. Only amusing stimulus ads were used while other emotional stimulus ads (e.g. disgusting or gloomy) are likely to evoke corresponding facial expressions. More research is in particular needed to test if negatively valenced stimuli can provoke facial expressions of happiness, as our low amusing video advertisements were at best neutral and not "negatively amusing". We did not include such stimuli because the within subject factor in the design could sustain carry-over effects between subsequent ads. In order to reduce priming effects any switches from extreme positively valenced stimuli to extreme negatively valenced ones had to be prevented.

As brought to our attention by an anonymous reviewer, correlations between facial expressions of happiness (FEH) and attitude toward ad (AAD) and brand (AB) are rather modest. For example, in high AVA, at best 37% of the variance in AAD is explained by the FEH-10% and as little as 7% in AB in medium AVA. There are many possible sources of variation in the scores of advertisement effectiveness. Brown and Stayman (1992) conducted a metanalysis of correlational pairwise relationships between AAD, AB and other variables. They found that AAD correlates significantly with AB, ad cognitions, brand cognitions, and purchase intentions. In addition, through structural path estimates, they demonstrated that ad cognitions explain 27 % of the variance in AAD, but AAD explains 32% of the variance in AB. They also found, as we did, that emotions explained around 25% of the variance in AAD. Therefore, we would argue that emotions (and expressions of emotions) and ad cognitions may jointly add to the formation of AAD, while in turn, AAD explains part of the variance in AB. However, we acknowledge, based on Eisend's (2011) review that other variables also are sources of the variance in both AAD and AB, including humor, brand- and ad- related positive and negative cognitive responses. In addition, it is likely that ad viewer characteristics too, such as gender, sex and socioeconomic status explain the variation in the scores.

Discussion

Facial expressions of happiness – automatically analyzed by FaceReader - can reliably distinguish between amusing and non-amusing video advertisements. In addition, at least in the amusing commercials it is likely possible to establish the advertisement effectiveness using facial expressions of happiness, given the fact that FaceReader measures correlated sufficiently high with participants' self-reports. We demonstrated that advertisement effectiveness of amusing video ads relates to the theoretically most obvious candidate emotion – happiness, and not with any other basic emotion.

Importantly, participants did not report they liked the advertisement because they were generally positively primed after watching an amusing stimulus. Facial expressions of happiness correlated with attitudes toward the advertisement and the brand only in the corresponding conditions. For example, there was no relation with how much people smiled in the high amusing condition and said they like the ad in a medium amusing condition. In addition, it could be the case that just novelty of the stimuli provoke people's positive attitude. We excluded that possibility by preselecting the stimuli that were amusing independent of previous experience with them. Importantly, we also showed that people did not regard the amusing stimuli as novel because they did not show facial expressions of surprise in any of the conditions. Our findings suggest that researchers should investigate further facial expressions that reflect emotions specific for an ad, like disgust expressions in response to disgusting ads, expressions of fear in response to scary ads, and so on. Our results lead one to expect that in each of these cases other basic emotions do not contribute to ad effects.

Moreover, we believe that the strength of our findings lay in the semi- field experiment set-up. We recorded people's reactions to persuasive stimuli in their natural environment - their houses and offices – i.e. the places where they usually watch such type of stimuli as ads. We

believe that the study of advertisement effectiveness should not be carried out only in laboratory settings but also should test theories in the real-life settings where actual advertisement consumption is taking place. In addition, we demonstrated that the available tools could reliably analyze the material collected under austere conditions. The video-recordings of the participants were of relatively poor quality. In many cases, the lighting conditions were suboptimal; the position of the participants' face was inadequate and their computers had too limited capacity to properly present the stimuli. The collected material permitted us to find significant differences and correlations in spite of an impoverished quality of data.

We acknowledge that there are other methods to analyze facial expressions, however those are primary focused on signal processing and are not exclusive to FaceReader (e.g. Teixeira, Wedel, & Pieters, 2012). Signal processing oriented analysis fits better with most data-intensive set-up studies such as the Teixeira et al., combining various real –time psychophysiological measures than with this study that focuses on relations of one such variable with a self-report criterion variable. The statistical method that we used helps to analyze rather noisy data and account for individual differences in a relatively simple way. People reacted to our stimuli with different facial expressions intensity and duration time. The top 10% peak values criterion gauges analyses to periods of intense reactions for each participant. The use of non-parametric tests for non-normally distributed data takes into account the skewedness of the participants' overall reactions, further accommodating the differences in facial reaction patterns. We believe that our approach constitutes another small step toward collection of reliable data in the field of affective consumer neuroscience.

However, further research is needed to overcome the limitations of the current study and discover more relations between patterns of facial expressions and people's reactions to persuasive stimuli. We recommend investigating the relation between facial expressions of

happiness and measures of advertisements effectiveness additional to the ones studied here. Purchasing intention and actual buying behavior seem obvious candidates. With current technology, it seems possible to setup a virtual mock supermarket where people see the presentation of the product (or actual advertisement) and can "buy" it. However, not only happiness expressions would predict such behavior but also facial expressions of other basic emotions. In our study, happiness related to the measures of advertisement effectiveness but these only covered attitudes; it seems probable that buying decisions are more affectively involving than ad and brand attitudes and evoking more complex patterns of facial expressions.

A final point we wish to make is that we believe we demonstrated that market researchers can potentially use FaceReader as a tool to measure advertisement effectiveness. We showed that people who like the advertisement more scored higher on one of the FaceReader measurements (i.e. happiness). We believe that we have contributed preliminary evidence validating FaceReader methodology for suitability in consumer research. The first step is to reliably assess differences in the facial reaction towards ads and relating these to self-reported effect-relevant scores. We set-up our study to test only effectiveness of the advertisements under amusing conditions, but further experiments could aim toward: (a) comparing laboratory and naturalistic studies of this kind to further demonstrate experimental validity; (b) testing different kind of persuasive stimuli (e.g. disgusting, scary or sad) and as said additional advertisement effectiveness measures; and (c) cross-validating FaceReader against other measures (e.g. AdSAM®, PrEmo, or facial EMG). We hope to address those and more questions in future studies.

References

Arnold, H. J., & Feldman, D. C. (1981). Social desirability response bias in self-report choice situations. *Academy of Management Journal*, *24*(2), 377-385.

Baumgartner, H., Sujan, M., & Padgett, D. (1997). Patterns of affective reactions to advertisements: The integration of moment-to-moment responses into overall judgments. *Journal of Marketing Research, 34*(2), 219-232.

Brown, S. P., & Stayman, D. M. (1992). Antecedents and consequences of attitude toward the ad: A meta-analysis. *Journal of Consumer Research, 19*(1), 34-51.

Buhrmester, M. D., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3-5. doi: 10.1177/1745691610393980

Chattopadhyay, A., & Basu, K. (1990). Humor in advertising: The moderating role of prior brand evaluation. *Journal of Marketing Research, 27*(4), 466-476.

Cole, P. M. (1986). Children's spontaneous control of facial expression. *Child Development, 57*, 1309-1321.

Cook, I. A., Warren, C., Pajot, S. K., Schairer, D., & Leuchter, A. F. (2011). Regional brain activation with advertising images. *Journal of Neuroscience, Psychology, and Economics, 4*(3), 147. doi: 10.1037/a0024809

Danner, L., Sidorkina, L., Joechl, M., & Duerrschmid, K. (2013). Make a face! Implicit and explicit measurement of facial expressions elicited by orange juices using face reading technology. *Food Quality and Preference*. 10.1016/j.foodqual.2013.01.004

de Wijk, R. A., Kooijman, V., Verhoeven, R. H. G., Holthuysen, N. T. E., & de Graaf, C. (2012). Autonomic nervous system responses on and facial expressions to the sight, smell, and taste of liked and disliked foods. *Food Quality and Preference, 26*(2), 196-203. doi: 10.1016/j.foodqual.2012.04.015

den Uyl, M., & van Kuilenberg, H. (2005). The FaceReader: Online facial expression recognition. In L.P.J.J. Noldus, F. Grieco, L.W.S. Loijens and P.H. Zimmerman (Eds.),

Proceedings of Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research (pp. 589-590). Wageningen, the Netherlands: Noldus Information Technology.

Derbaix, C. M. (1995). The impact of affective reactions on attitudes toward the advertisement and the brand: A step toward ecological validity. *Journal of Marketing Research, 32*(4), 470-479.

Desmet, P. M. A. (2002). *Designing emotions* (Doctoral dissertation). Retrieved from http://resolver.tudelft.nl/uuid:937aaf05-12f0-4283-adfb-007bdbbeaaa2

Eisend, M. (2011). How humor in advertising works: A meta-analytic test of alternative models. *Marketing Letters, 22*(2), 115-132. doi: 10.1007/s11002-010-9116-z

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation, 1971* (vol. 19, pp. 207–282). Lincoln: University of Nebraska Press.

Ekman, P., & Cordano, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3*(4), 364-370. doi: 10.1177/1754073911410740

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen W.V., & Hager J. C. (2002). *Facial action coding system: The manual*. Salt Lake City, UT: Research Nexus.

Friesen, W. V., & Ekman, P. (1983). *EMFACS-7: Emotional facial action coding system*. San Francisco, CA: University of California, San Francisco.

Greenwald, M. K., Cook, E. W., & Lang, P. J. (1989). Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology 3*, 51-64.

Gross, J.J., & John, O.P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348-362. doi: 10.1037/0022-3514.85.2.348

Gross, J.J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J.J. Gross (Ed.), *Handbook of emotion regulation*. New York: Guilford Press.

Havlena, W. J., & Holbrook, M. B. (1986). The varieties of consumption experience: Comparing two typologies of emotion in consumer behavior. *Journal of Consumer Research, 13*(3), 394-404.

He, W., Boesveldt, S., de Graaf, C., & de Wijk, R.A. (2012). Behavioural and physiological responses to two food odours. *Appetite*, 59 (2), 628.

Hopkins, R., & Fletcher, J. E. (1994). Electrodermal measurement: Particularly effective for forecasting message influence on sales appeal. In A. Lang (Ed.), *Measuring psychological responses to media messages* (pp. 113-132). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lajante, M., Droulers, O., Dondaine, T., & Amarantini, D. (2012). Opening the "black box" of electrodermal activity in consumer neuroscience research. *Journal of Neuroscience, Psychology, and Economics*, *5*(4), 238-249. doi: 10.1037/a0030680

Langleben, D. D., Loughead, J. W., Ruparel, K., Hakun, J. G., Busch-Winokur, S., Holloway, M. B., ... & Lerman, C. (2009). Reduced prefrontal and temporal processing and recall of high "sensation value" ads. *Neuroimage*, *46*(1), 219-225. doi: 10.1016/j.neuroimage.2008.12.062

Meyers-Levy, J., & Malaviya, P. (1999). Consumers' processing of persuasive advertisements: An integrative framework of persuasion theories. *Journal of Marketing, 63*, 45-60.

Micu, A. C., & Plummer, J. T. (2010). Measurable emotions: How television ads really work patterns of reactions to commercials can demonstrate advertising effectiveness. *Journal of Advertising Research, 50*(2), 137-153. doi: 10.2501/S0021849910091300

Mitchell, A., & Olson, J. C. (1982) Are product attribute beliefs the only mediator of advertising effects on brand attitude? *Journal of Marketing Research, 18*, 318-332.

Morris, J. D., Chongmoo, W., Geason, J. A., & Jooyoung, K. (2002). The power of affect: Predicting intention. *Journal of Advertising Research, 42*(3), 7-17.

Noldus. (2013) FaceReader: Tool for automatic analysis of facial expression: Version 5.0.15. Wageningen, the Netherlands: Noldus Information Technology B.V.

Ohme, R., Reykowska, D., Wiener, D., & Choromanska, A. (2009). Analysis of neurophysiological reactions to advertising stimuli by means of EEG and galvanic skin response measures. *Journal of Neuroscience, Psychology, and Economics*, *2*(1), 21-31. doi: 10.1037/a0015462

Olney, T. J., Holbrook, M. B., & Batra, R. (1991). Consumer responses to advertising: The effects of ad content, emotions, and attitude toward the ad on viewing time. *Journal of Consumer Research, 17*(4), 440-453.

Phillips, B., J. (2000). The impact of verbal anchoring on consumer response to image ads. *Journal of Advertising, 29*(1), 15-24.

Pieters, R., & Wedel, M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing, 68*(2), 36–50.

Poels, K., & Dewitte, S. (2006). How to capture the heart? Reviewing 20 years of emotion measurement in advertising. *Journal of Advertising Research, 46*(1), 18–37. doi: 10.2501/S0021849906060041

Pryor, J. B., Gibbons, F. X., Wicklund, R. A., Fazio, R. H., & Hood, R. (1977). Self-focused attention and self-report validity. *Journal of Personality*, *45*(4), 513-527.

Ramsøy, T. Z., Friis-Olivarius, M., Jacobsen, C., Jensen, S. B., & Skov, M. (2012). Effects of perceptual uncertainty on arousal and preference across different visual domains. *Journal of Neuroscience, Psychology, and Economics*, *5*(4), 212-226. doi: 10.1037/a0030198

Ravaja, N. (2004). Contributions of psychophysiology to media research: Review and recommendations. *Media Psychology*, *6*(2), 193-235. doi: 10.1207/s1532785xmep0602_4

Reisenzein, R., Studtmann, M., & Horstmann, G. (2013). Coherence between emotion and facial expression: Evidence from laboratory experiments. *Emotion Review*, *5*(1), 16–23. doi: 10.1177/1754073912457228

Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-induced engagement in internet video advertisements. *Journal of Marketing Research*, *49*(2), 144-159. doi: 10.1509/jmr.10.0207

Terzis, V., Moridis, C. N., & Economides, A. A. (2012). The effect of emotional feedback on behavioral intention to use computer based assessment. *Computers & Education*, *59*(2), 710-721. doi: 10.1016/j.compedu.2012.03.003

Terzis, V., Moridis, C. N., & Economides, A. A. (2013). Measuring instant emotions based on facial expressions during computer-based assessment. *Personal and Ubiquitous Computing*, *17*(1), 43-52. doi: 10.1007/s00779-011-0477-y

van Kuilenburg, H., Wiering, M., den Uyl, M. (2005). A model based method for facial expression recognition. In D. Hutchison, T. Kanade, J. Kittler, J.M. Kleinberg, F. Matern, J.C. Mitchell, M. Noar, … G. Weikum (Eds.), *Lectures Notes in Computer Science: Vol. 3720. Machine Learning: ECML 2005* (pp. 194-205). Berlin, Germany: Springer-Verlag. doi: 10.1007/11564096_22

Wedel, M., & Pieters, R. (2000). Eye fixations on advertisements and memory for brands: A model and findings. *Marketing Science, 19*(4), 297–312. doi: 10.1287/mksc.19.4.297.11794

Wiles, J. A., & Cornwell, T. B. (1990). A review of methods utilized in measuring affect, feelings, and emotion in advertising. *Current Issues & Research in Advertising*, *13*, 241–275.

Winkielman, P., Berntson, G. G., & Cacioppo, J. T. (2001). The psychophysiological perspective on the social mind. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: intraindividual processes* (pp. 89–108). Oxford, England: Blackwell.

Zeitlin, D. M., & Westwood, R. A. (1986). Measuring emotional response. *Journal of Advertising Research*, *26*(5), 34-44.