Journal of Marketing Research Article Postprint

EMOTION-INDUCED ENGAGEMENT IN INTERNET VIDEO ADS

Thales Teixeira[1], Michel Wedel,[2] and Rik Pieters[3]

- Forthcoming at *Journal of Marketing Research* -

1 Assistant Professor, Harvard Business School, Harvard University, Boston, MA 02163, USA, Tel. (617) 495-6125, tteixeira@hbs.edu.

2 Professor of Marketing, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA, Tel. (301) 405-2162, Fax. (301) 405-0146, mwedel@rhsmith.umd.edu.

3 Professor of Marketing, Tilburg University, Faculty of Economics and Business, Department of Marketing, P.O. Box 90153, 5000 LE Tilburg, the Netherlands, Tel: + 31 13 466 32 56, Fax. + 31 13 466 28 75, f.g.m.pieters@uvt.nl.

The authors would like to thank Nicu Sebe (University of Trento) for providing the facial expressions classification algorithm, and Cherry Kwun for help with the experiment. The first author would like to thank John Gourville for his valuable suggestions regarding an earlier version of this paper. The usual disclaimer regarding responsibility for mistakes applies.

EMOTION-INDUCED ENGAGEMENT IN INTERNET VIDEO ADS

ABSTRACT

This study shows how advertisers can leverage emotion and attention to engage consumers in watching Internet video ads. In a controlled experiment, joy and surprise were assessed through automated facial expression detection for a sample of ads. Concentration of attention was assessed through eye tracking, and retention of viewers by recording zapping behavior. This allows tests of predictions about the interplay of these emotions and inter-individual attention differences at each point in time during exposure. Surprise and joy effectively concentrate attention and retain viewers. But importantly, the level rather than the velocity of surprise impact attention concentration most, whereas the velocity rather than the level of joy impact viewer retention most. The effect of joy is asymmetric, with higher gains for increases than losses for decreases. Based on these findings, we develop representative emotion trajectories to support ad design and testing.

Keywords: Internet video avoidance, emotions, attention, eye tracking, facial expressions, simultaneous dynamic model, frailty model.

Ad avoidance is a major concern for advertisers and broadcasters. Certain demographics avoid up to 80% of television commercials that they are exposed to and 87% of digital video recorder owners often actively skip past ads (Grover and Fine 2006). Internet video advertising is therefore increasingly seen as an opportunity. With more and more video ads migrating to the web, they now appear as in-page video ads, or before, during, or after streaming, animation, or gaming content (Elkin 2010). Yet, along with a steady growth to over $700 million in 2008, concern over avoidance of these video ads has grown as well[1]. In terms of industry economics, when targeted viewers zap, skip, zip or click past video ads, the firm's brand loses the opportunity to communicate, the broadcaster loses viewers, and the website loses exposure.

Evoking positive emotional responses is considered a potent strategy to engage consumers from moment-to-moment in video ads, that is, to attract their attention and retain them from start to finish. Such emotion-induced engagement of viewers increases the likelihood of obtaining desired down-stream communication effects (Vakratsas and Ambler 1999). But how does one engage consumers emotionally, and how do the resulting emotions influence avoidance decisions from moment to moment in video ads? These questions motivated the current study.

Psychology has made great strides in understanding attention effects of negative emotions, such as fear and anxiety (Yiend 2010), but less is known about the positive emotions that are dominant in consumer advertising. Advertising research in its turn has focused on emotions such as joy and surprise, "…to engineer positive environments for consumers" (Griskevicius et al. 2010, p. 238). Yet, it has mostly examined down-stream rather than immediate effects such as zapping, and it has emphasized valence and activation dimensions rather than specific emotions such as joy and surprise (Baumgartner, Sujan, and Padgett 1997; Olney, Holbrook, and Batra 1991). The dynamic effects that specific positive emotions have on

[1] http://www.iab.net/media/file/IAB_PwC_2008_full_year.pdf, last accessed May 2010.

consumer engagement with advertising are thus still largely unexplored. Moreover, research about the moment-to-moment interplay between emotions and attention is scarce both in advertising and psychology (Fredrickson and Branigan 2005; Yiend 2010).

Studying attention and the emotions that consumers experience from moment-to-moment may reveal the potential of joint emotion and attention tracking in pretesting advertising and indicate how advertisers can engage viewers in video and television ads. A recent study has found that television ads that concentrate viewers' visual attention in a small region of the screen, from moment-to-moment, rather than allow their attention to freely wander and disperse across the screen, retained viewers longer and were zapped less (Teixeira, Wedel, and Pieters 2010). However, that study did not investigate the antecedents of attention concentration, and specifically if and how the emotions evoked by ads accomplish this. This is the focus of the present study. It examines the influence that joy and surprise, two related emotions, have on moment-to-moment engagement of consumers in video ads. These two emotions are among the most commonly targeted emotions in advertising, as prior studies (Derbaix 1995) and a content analysis of on-line video ads reveal. For that content analysis, ten trained coders identified the targeted emotions (intended to be evoked) in a sample of 106 randomly selected on-line video ads (July 2010). Joy/Happy was targeted most frequently (38%), followed by surprise (26%), and lagged by disgust (6%), with the other emotions (anger, fear, sadness) trailing even more, and 18% of ads having no clear target emotion.

We use moment-to-moment measures of emotions and attention evoked by video ads on a website to predict the dynamics of consumers' avoidance decisions. We assess joy and surprise via automated facial expression detection, attention concentration via infrared tracking of eye-movements, and viewer retention by recording consumers' decisions to continue watching or zap the ads. We let viewers freely decide what to watch and what not, by zapping a video ad at any point in time. This self-controlled exposure enables emotional reactions to have the same behavioral effects as in real-life situations, which is of obvious importance for diagnostic purposes (Derbaix 1995). We collect data on emotions, attention, and zapping behavior at 250 millisecond intervals, across 28 video ads for 50 viewers. This results in 145,000 frames of data, which we augment with control variables known to independently influence zapping. We test our predictions using a simultaneous Bayesian Frailty model, estimated with MCMC, accounting for observed and unobserved temporal, individual and stimulus sources of heterogeneity.

Companies and research companies, such as Procter & Gamble, Unilever and GfK, have recently started to collect high frequency data on emotions from facial expressions to understand their influence of consumer behavior. This has been enabled by software developments that automate collection of this unique type of data, such as by eMotion (University of Amsterdam), FaceReader (Noldus) and OKAO (Omron Corp). A challenge has been, however, to develop appropriate methods that can extract diagnostic information from the resulting massive and noisy data streams and relate this information to the dependent variables of interest, such as zapping. To the best of our knowledge, our study is the first to develop such a statistical approach to eye and facial expression tracking in advertising.

EMOTIONAL ENGAGEMENT WITH VIDEO ADS

People experience emotions when their personal interests are at stake. Over the course of evolution, a set of basic emotions has developed each with their own eliciting conditions, experiential content, facial expressions and behavioral tendencies (Plutchik 1980). Facial expressions serve to communicate emotions to both self and others. Distinct and cross-culturally universal facial expressions have been found, amongst others, for the emotions of joy, surprise, sadness, disgust, anger and fear (Ekman 1999). Facial expressions of emotions, such as joy and surprise, differ and can be measured continuously and non-intrusively (Derbaix 1995; Wehrle et al. 2000). This offers advantages over self-report measures that may lack this sensitivity, are slow and difficult to assess continuously, and may lead to mere measurement effects, which all threaten their validity.

Ever since Darwin (1872) reported on the strong link between experienced emotions and facial expressions, there has been a keen interest in developing methodologies to accurately and efficiently assess them. The Facial Actions Coding System (FACS; Ekman and Friesen 1978) to identify basic emotions from facial expressions has proven useful in marketing contexts, but has relied on manual coding of video footage (Lemmink and Mattsson 1998). For example, Derbaix (1995) used ten human coders to measure FACS reactions to 13 ads in intervals of one second. Yet, manual coding is error-prone, laborious, and difficult at the high frequencies at which emotions such as joy and surprise unfold. This has prevented its wide-scale use in marketing.

For these reasons, we use computer-aided emotion detection from facial expressions. Bartlett et al. (1999) early on showed that these detection algorithms outperformed non-expert coders and were about as accurate as expert coders. Moreover, computer-aided emotion detection can now be done, in real time, at a rate of 4Hz (every 250 milliseconds), which is much faster than achievable by human coders (Cohen et al. 2003). It provides the temporal resolution needed to identify the fast acting effects of emotions on visual attention and behaviors such as zapping, which can occur well within one second (Nummenmaa et al. 2009).

Dynamics of Emotions

We build on the “modal model” of emotion regulation (Gross and Thomson 2007) to understand the dynamics of the emotional response to video commercials (shown schematically in Figure 1). Emotion regulation refers to the processes by which individuals regulate either positive or negative emotions over time, in either an automatic, unconscious or controlled, conscious manner. The emotional responses generated by the appraisal of the video commercial include fast acting facial expressions.

[Insert figure 1 about here]

The purpose of emotion regulation is to reduce, maintain or intensify the emotional experience resulting from exposure to the stimulus. The modal model specifies a recursive relation between the stimulus and three key components: emotion, attention, and behavior. While there are several ways to regulate emotions, we focus here on attention deployment and stimulus selection. Attention deployment refers to how individuals direct their attention to a stimulus in order to regulate their emotions. Distraction and concentration are two major attention strategies in response to emotionally salient aspects of the stimulus (Gross and Thomson 2007). Whereas concentration focuses attention to these emotionally salient aspects, distraction directs attention away from them. Stimulus selection is the most forward-looking approach, and involves approach/avoidance reactions to the emotion-eliciting stimulus, such as when watching/zapping a video commercial. Emotion regulation is a dynamic process, involving feedback from the emotion to attention and behavioral responses to the stimulus, as shown in Figure 1. Thus, exposure to a stimulus elicits an emotion. To the extent that the emotion is shared by individuals, they will similarly focus attention to, or away from, and engage, or avoid, the stimulus in order to abate, maintain or intensify the emotion. This generates a new emotional experience, which needs to be abated, maintained or intensified, and so on. Avoidance (zapping) ends the stream of emotional stimulation. This dynamic regulation process implies that emotions affect attention and both will affect zapping, the magnitude of these effects changing during exposure to a video ad. Our dynamic statistical model, described in the Model section, captures such effects.

Effects of Joy and Surprise

Emotions serve to organize perception and action in order to attain specific goals (Plutchik 1980). Negative emotions prompt tendencies to avoid or reject the affective stimulus, and positive emotions tendencies to approach or retain it. Surprise and joy are the focus of the current study and we derive predictions about their effects on consumer engagement with ads.

Joy and other positive emotions have been widely recognized to activate tendencies to broaden attention (Frederickson 1998). As a case in point, Fredrickson and Branigan (2005) found that positive emotions induced through short film clips, led to more visual exploration as measured via self-report and facial muscle electromyography. However, recently it has been recognized that this broadening effect of positive emotions only holds once people have attained their current goals, in low motivation conditions. Positive emotions then spur consideration of other (internal) goals or (external) stimuli in the environment.

However, before current goals have been attained positive emotions such as joy have been shown to increase attentional focus in high approach motivation conditions to further successful goal pursuit (Gable and Harmon-Jones 2008). According to emotion regulation theory (Gross and Thomson 2007; Figure 1), such conditions would involve both attention concentration and stimulus selection as key regulation strategies. We believe that exposure to video ads presents such a condition. The experience of a positive emotion in a video ad informs the person that ad is beneficial and activates a goal to continue exposure (Plutchik 1980). This goal encourages an increase in attentional focus, which assists in attaining the goal and promotes action tendencies to continue or maintain goal pursuit (Gable and Harmon-Jones 2008, p.481).

During exposure to video ads, joy will thus not only increase attentional focus, but also induce action tendencies to continue watching the ad. This is consistent with the prior finding that the moment-to-moment pleasantness of television ads reduces zapping (Woltman-Elpers, Wedel, and Pieters 2003). For emotional pictures, related effects on psychophysiological measures (Hajcak and Olvet 2008) and eye movements (Calvo and Lang 2004) have been reported.

Eye-movements on ads comprise of sequences of fixations--when the eye is still and information is acquired--and saccades--fast jumps to direct the focus of attention to another area of interest from which information is acquired. Increased attentional focus would express itself in fixation patterns that are less variable across individuals, because idiosyncratic differences in interest are reduced so that individuals tend to focus attention similarly. This inter-individual concentration of attention is not to be confounded with intra-individual concentration, the concept measured in most prior research (e.g., Germeys and d'Ydewalle 2007; Gable and Harmon-Jones 2008). The basic idea is that when people are following the ad as intended by its designer, most will look at roughly the same thing on the screen. Individuals who are losing interest will be distracted by other aspects of the ad and therefore will differ in their fixation location from the rest, at that point in time. One would thus predict that the emotion of joy leads to more focused attention patterns, i.e. individuals looking more at the point of focal interest in the ad, at each point in time. We also expect that, independent of its visual attention effects, joy stimulates viewer retention by reducing the probability to stop watching video ads.

Surprise arises when outcomes are unexpected. It is characterized as hedonically neutral or as positive or negative, and related to feelings of interest, curiosity, wonder and joy (Frederickson 1998). The latter are all positive emotions. Surprise informs the person that prior expectations are disconfirmed. This leads to interruption of ongoing information processing and reorientation to the possibly significant event (Meyer, Reisenzein and Schutzwohl 1997). Surprise developed to grab people's attention fast and motivate them to engage in specific action (Plutchik 1980). Thus, we expect surprise to stimulate orientation to the source of unexpectedness, which would be reflected in patterns of attention across participants that are concentrated on the source, rather than dispersed. Because of its vigilant nature, we expect the concentrating effect of surprise on attention to be stronger that of joy.

Because it requires time to resolve the expectation disconfirmation and raises interest, surprise should stimulate viewer retention as well. However, because the experience of surprise often is hedonically neutral or negative, the urge to prolong exposure falls rapidly once expectation disconfirmation is resolved. Therefore, we expect the effect of surprise on viewer retention to be less strong than the effect of joy. In sum, we test:

H1: Joy and surprise increase the concentration of attention across viewers on the same visual locations in video ads (a), with the effects of surprise being stronger (b).

H2: Joy and surprise increase viewer retention in video ads (a), with the effects of joy being stronger (b).

Attention concentration (dispersion) is the extent to which a viewer focuses (diverges) attention to (from) a single location, *at each point in time*. The assumption is that at each time-frame in the ad, there is a single location of intended focus and that this location corresponds with the consensus region that participants fixate on. Over time the intended locus of attention may shift, or new ones may appear. Thus, our hypotheses make predictions about the extent to which the individual fixation points conform to those of the crowd, at each point in time, reflecting focus of attention. The relevant comparison is thus across people within each time-frame, because the intended attentional locus changes over time. Both H1 and H2 refer to the downstream effects of the felt emotions on the individual attentional focus.

We predict that zapping will be less when ads minimize heterogeneity among individuals in regard to what they focus on, irrespective of what the object is, that is, when they "bind" attention. If the findings of Teixeira, Wedel, and Pieters (2010) hold up for the present study as well, attention concentration prolongs viewer retention by decreasing zapping. Then, support for both hypotheses would imply that surprise improves viewer retention directly by weakly decreasing zapping probabilities, and indirectly by strongly concentrating attention, which in turn also decreases zapping probabilities. It would also imply that joy improves viewer retention directly by strongly decreasing zapping, and that it indirectly improves viewer retention by weakly concentrating attention, which decreases zapping. Thus, joy and surprise would both improve engagement with video ads directly and indirectly, but in very different ways. To test these predictions, we examine both the level and the velocity of each emotion as separate features of their moment-to-moment traces during ad exposure. The level of an emotion is its intensity at a given moment during ad exposure. The velocity (change) of an emotion is indicated by the first-order derivative of the emotion trace. There is evidence that both measures are important from studies on satisfaction with gambles (Hsee and Abelson 1991), on influence of moment-to-moment overall feelings on ad evaluations (Baumgartner, Sujan and Padgett 1997) and on how the pleasantness of television ads influences zapping (Woltman-Elpers, Wedel and Pieters 2003). It is therefore reasonable to predict that the level and velocity of joy and surprise separately influence moment-to-moment attention and zapping decisions. But because this is not a previously studied topic, we cannot make specific predictions about the relative magnitude of the level versus velocity effects. Yet, we think that empirical findings on velocity (change) of joy and surprise may provide key insights and help further theorizing.

DATA

Participants and Stimuli

Fifty-eight paid students and staff members (mean age 22, range 18 to 49 years, 53% male) of a major Northeastern American university participated in a controlled experiment on on-line browsing behavior. Participants were in the target audience for the video ads.

Participants were exposed to 28 video ads, 14 emotional and 14 neutral, in an online setting. The neutral ads were interspersed between emotional ads as buffers to reduce the mental load on participants. Emotional ads were expected to evoke joy or surprise at some point in time (based on a pretest with 14 other participants and 21 video ads, from which the 14 target ads were chosen). Ads were individually embedded in identical web pages in the form of post-rolls (sequence of ads, each of which automatically played upon page loading after the video content). Half of the target ads were selected to evoke joy (e.g., smile, laughter) or surprise (e.g., elevation of eyebrows, mouth open), or both, as confirmed by the pretest. The final list of ads chosen is provided in Table 1, with the average emotion intensity across time and participants, the average intensity of positive surprise and the number of participants that expressed each emotion for at least one second. The ad elements that evoked emotions in our data vary, from jokes to specific images or scenes and are not the focus of this research. Ads were for various categories (e.g., beverages, consumer packaged goods, telecom, cleaning supplies and financial services) for well-known (e.g., Budweiser, Nivea, Dell) and lesser-known brands (e.g., Lincoln Insurance, Mercator, Rockstar Drinks). The ad sequence was counterbalanced to reduce order effects. Because ads were not randomly chosen, our analysis aims to discover emotion's potential, not its estimate based on a representative sample of ads in the marketplace.

[Insert table 1 about here]

Data Collection

Participants' visual attention, facial emotion expressions and zapping decisions during exposure to video ads were simultaneously assessed. Participants were seated in a quiet room in front of a 17-inch eye-tracker monitor, with a separate camera affixed to the top of the monitor for facial expression recording. To relax them and increase realism,[2] participants first saw a short four-minute humorous sitcom clip followed by the video ads. Participants were instructed that they could either watch each ad until the end, zap to the next one by pressing the space bar or click on the link provided at the bottom of the page with a mouse to go to the advertised brand's web page. While the web pages containing ads were identical, apart from the videos themselves, a click on the link sent participants to a one-page mock-up site specific for that brand, and provided additional brand information, but not so much that participants would stay too long at the brand's site. The link at the end of this webpage brought them to the next video ad in the sequence. Participants were asked to keep one hand over the space bar and the other over the mouse at all times. Showing ads on web pages with links increases realism and supports that the study was indeed on online browsing behavior, maximizing the opportunity for regular facial expressions to show up. No further analysis was done to understand clicking behavior. At the end all participants were informed of the true aims of the experiment.

A Tobii 1750 infrared eye-tracker unobtrusively (no head or chin gear) measured eye movements using infrared cameras at the rate of 50 Hz with spatial resolution of less than 0.5 degrees of visual angle. Participants had complete freedom of head movement.

Facial expression footage from each participant was collected by means of a MiniDV camera coupled to the eye-tracker and aimed at the participant's face. The continuous video images served as input to the emotion detection software, which works by fitting a virtual

[2] The low average reported levels of stress (1.43), nervousness (1.51), feeling of being observed (1.51) and abnormal viewing behavior (1.60) on a 5 point scale, support the success of this.

facemask to the video image of the face. This facemask adjusts to the form of the face (eyes, eyebrows, nose, face and mouth delimiters) so as to capture 64 deviations in the line segments that relate to Ekman's FACS. These measures were processed online at the rate of 4 Hz using a Bayesian Neural Network Classifier calibrated on the images of the Cohn-Kanade database, a well-known benchmark (Cohen et al. 2003) with 500 images from 100 people. If a participant smiles, for example, some of the deviations in line segments will increase, such as the one linking both corners of the lips, while others will decrease, such as the ones linking corners of the lips to the cheekbones or to the eyes. The output of the classifier is the probability that the viewer exhibits the emotion, or a neutral state. We use the probability measures for joy and surprise. Hit rates of respectively 86% and 94% were assessed using cross-validation for a test subset of the Cohn-Kanade database (Cohen et al. 2003). After participants had been exposed to the complete reel of ads, each participant was taken to a computer in another room to answer questions about the ads and themselves. The experiment lasted about 45 minutes.

Measures

Table 2 provides summary statistics of measures to be described next.

Ad avoidance. The criterion is the instant of the zapping decision, if taken, i.e., when a participant stops watching a particular commercial by pushing the space bar (1 = avoid, 0 = else). Since this event can only occur once for a participant-ad combination, zapping at a time frame is always conditional on not zapping previously. This dependent variable represents a binary cross-sectional (participants) unbalanced repeated measures (ads) time-series of zapping decisions.

Emotions. The output of the emotions detection algorithm is a classification accuracy measure, ranging from 0 to 1 for each time frame for joy and surprise. Higher values indicate a higher likelihood that a viewer experiences the respective emotion at each 250 msec. instant. Since accuracy is related to intensity, this measure also serves as a proxy for emotional intensity.

Attention dispersion. Attention is concentrated when for a particular frame in the ad, multiple eye fixations of participants cluster on a small spatial region in the ad, reflecting a consensus region of attentional focus, and it is dispersed when multiple eye fixations land on a large spatial region. Following Teixeira, Wedel, and Pieters (2010), we calculate two measures of (lack of) attention concentration: Individual (IAD) and Aggregate Attention Dispersion (AAD) using the X and Y-coordinates of focal eye positions detected by the eye-tracker. As a function of the X and Y focal position vector, f_{ict}, for individual i, ad c and time-frame t, IAD and AAD (for all N_t individuals that have not zapped at t) are:

$$\mathrm{IAD}_{ict} = \left(f_{ict} - \tfrac{1}{N_t} \sum_{i=1}^{N_t} f_{ict} \right)' \left(f_{ict} - \tfrac{1}{N_t} \sum_{i=1}^{N_t} f_{ict} \right) \qquad \mathrm{AAD}_{ct} = \frac{1}{N_t} \sum_{i=1}^{N_t} IAD_{ict} \, . \tag{1}$$

IAD is calculated as a squared Euclidian distance from the moment-to-moment centroid, which is computed as the average of the fixation locations of all individuals, for each ad c at each instant t. IAD measures inter-individual attention concentration (smaller values reflecting higher concentration), and AAD is used as a control variable. IAD is a relative (to the average focal point across all individuals at each point in time) measure of individual attention dispersion. It has the desirable property of not requiring any content analysis of specific locations and thus can be used at high temporal resolutions. Lower attention concentration expresses lack of momentary ad engagement. We included the interaction between IAD and AAD to capture the effect of attention concentration by an individual when most other viewers have dispersed attention.

Control variables. We controlled for other characteristics of video ads and viewers that may influence attention and zapping decisions independent of emotions (Teixeira, Wedel, and

Pieters 2010; Woltman-Elpers, Wedel, and Pieters 2003), namely branding (presence of logo, brand name) and visual complexity of all video ad frames, and age, gender and familiarity of the participants with the video ad and brand (Table 2).

[Insert table 2 about here]

Data Aggregation and Development of Measures

Eye movement, facial expression and stimulus frame data were aggregated to a 250 msec. time frame. This time frame is within the average fixation duration for dynamic stimuli, and presents a lower bound for visually based response latencies (Mihaylova et al. 1999). We used a two-step procedure to determine the key measures in the time-courses of the two emotional expressions. First, we applied Functional Data Analysis (FDA; Ramsay, Hooker, and Graves 2009) using the S+FDA package (Clarkson et al. 2005) to identify instances of joy and surprise from the raw software output, while controlling for individual differences and measurement error. We choose 3rd degree polynomial B-splines as basis functions and set the FDA smoothing parameter to be equal to λ=102 (Ramsay, Hooker, and Graves 2009; Hsee and Abelson 1991). Second, the functional curves of viewers were clustered separately for each ad and type of emotion using Wards' method with Tibshirani et al.'s (2001) Gap procedure, which provides an optimal number of clusters (see Web Appendix A). The resulting traces were used to determine the level and velocity (1st derivative) of joy and surprise at each time frame during ad exposure[3]. Because positive velocities in joy had different effects than negative ones, we obtain both the velocities and absolute velocities. Previous research has shown some evidence for emotional change asymmetries in the context of advertising (Olsen and Pracejus 2004)[4].

---

[3] We also tested velocity changes (acceleration) but found no evidence for cross-cluster differences of this measure.
[4] Future work could unify the smoothing and clustering steps with the model estimation into a single, and more elegant, method. We thank one of the reviewers for pointing this out.

MODEL

We develop a model for the moment-to-moment zapping decisions and attention concentration, linking them to the emotions and control variables as specified by the modal model of emotional regulation (Gross and Thomson 2007). Our model is a bivariate mixed-outcome dynamic frailty model. We assume that the probability that individual i decides to avoid ad c at time-frame t, given parameters Θ_t, is $P(y_{ict} = 1 \mid \Theta_t) = \pi_{ict}$, where $y_{ict} = 1$ if participant i zaps commercial c at time t, and $y_{ict} = 0$ otherwise. We assume additive separability of strictly individual, commercial and time-specific baseline avoidance rates. Models with additive separable individual and trial (here ads) random effects are known as Frailty Models and are widely used due to their parsimony. In our study, this formulation causes a reduction of the *ict*-specific fixed and random effects parameters from over 300,000 to 318. Our model allows emotions to have time-dependent effects on decisions such as zapping (Kahneman 2000).

We formulate a binary probit duration model (Sueyoshi 1995), with the exogenous influence of individual and stimuli-specific regressors and time-varying coefficients as:

$$\Phi^{-1}(\pi_{ict}) = \mu_i^{(1)} + \alpha_c^{(1)} + X_{ict}^{(1)}\theta_t^{(1)} + Z_{ict}^{(1)}\Psi^{(1)} \quad (2)$$

$\Phi(\cdot)$ is the Normal cdf. The right-hand side (RHS) of the zapping equation contains the individual, ad and temporal baseline zapping rates, followed by the aggregate effects, linearly associated to the expectation of the dependent variable through the Probit link. $X^{(1)}$ is made up of the emotion variables, allowed to have time-varying effects θ_t. $Z^{(1)}$ includes attention dispersion, visual complexity, and brand placement covariates. Due to collinearity between some of the latter, we incorporate only brand (logo or name) Presence, its Duration and Cardinality of the order of appearance. Also, as detailed in the next section, we used lagged effects of some of the variables to accommodate delays in consumer responses and the dynamics of emotion regulation.

To measure the extent to which the emotions affect moment-to-moment ad zapping decisions, above and beyond the direct effect in Equation 2, we need to estimate the indirect, mediating role of attention dispersion as well, see Figure 1. Given that IAD is a squared Euclidian distance (positive and skewed), we model it with a log-linear Frailty model:

$$\mathrm{Log}(\mathrm{IAD}_{\mathrm{ict}}) = \mu_{\mathrm{i}}^{(2)} + \alpha_{\mathrm{c}}^{(2)} + \theta_{\mathrm{t}}^{(2)} + \mathrm{X}_{\mathrm{ict}}^{(2)}\Psi^{(2)} + \varepsilon_{\mathrm{ict}}^{(2)} \ , \tag{3}$$

where, similar to the zapping model, the RHS incorporates demographics, individual-ad and $X^{(2)}$ captures the emotions and other covariates as before.

In order to incorporate the covariation in zapping and IAD, we jointly estimate the effect of emotions, attention and control variables on zapping, and the effect of emotions and control variables on attention, see Figure 1. Stacking the error terms into ε_{ict}, and letting ${}^{*}Z^{(1)}$ be $Z^{(1)}$ less the attention dispersion variables, the simultaneous model can be written as:

$$\begin{pmatrix} \mathrm{Probit}(\pi_{ict}) \\ \log(IAD_{ict}) \end{pmatrix} = \begin{pmatrix} \mu_i^{(1)} + \alpha_c^{(1)} + X_{ict}^{(1)}\theta_t^{(1)} + {}^{*}Z_{ict}^{(1)}\Psi^{(1)} + \psi_1 IAD_{ict} + \psi_2 AAD_{ct} + \psi_3 (\mathrm{IAD}_{ict} \times \mathrm{AAD}_{ct}) \\ \mu_i^{(2)} + \alpha_c^{(2)} + \theta_t^{(2)} + X_{ict}^{(2)}\Psi^{(2)} \end{pmatrix} + \boldsymbol{\varepsilon}_{ict}, \tag{4}$$

with:

$$\begin{pmatrix} \mu_i^{(1)} \\ \mu_i^{(2)} \end{pmatrix} \sim \begin{matrix} N(X_i^{(3)}\Lambda^{(1)}, V_\lambda^{(1)}) \\ N(X_i^{(3)}\Lambda^{(2)}, V_\lambda^{(2)}) \end{matrix} \qquad \begin{pmatrix} \theta_t^{(1)} \\ \theta_t^{(2)} \end{pmatrix} \sim \begin{matrix} N(\theta_{t-1}^{(1)}, \omega^{(1)}) \\ N(\theta_{t-1}^{(2)}, \omega^{(2)}) \end{matrix}$$

$$\begin{pmatrix} \alpha_c^{(1)} \\ \alpha_c^{(2)} \end{pmatrix} \sim \begin{matrix} N(X_c^{(4)}\mathrm{K}^{(1)}, V_\kappa^{(1)}) \\ N(X_c^{(4)}\mathrm{K}^{(2)}, V_\kappa^{(2)}) \end{matrix} \qquad \varepsilon_{ict} \sim N\left(0, \begin{bmatrix} 1 & \sigma_{12}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 \end{bmatrix}\right)$$

Equation 4 describes a bivariate mixed outcome dynamic frailty model. μ_i and α_c are individual and ad specific, respectively, and are a linear function of demographics ($X_i^{(3)}$ containing age and gender, and with associated parameters Λ) and individual-ad characteristics ($X_c^{(4)}$ containing ad duration, ad and brand familiarity, with associated parameters K). Moreover, we allow the time-specific baselines θ_t to evolve stochastically according to a random walk to capture non-monotonic behaviors (Gustafson and Siddarth 2007).

In sum, the Probit model describes zapping on a moment-to-moment basis and as a direct and indirect function of emotions, as well as attention measures and other covariates that capture individual, stimulus and temporal heterogeneity. The Log-Linear model does the same for attention concentration (IAD). The joint model provides a dynamic representation of zapping and attention consistent with Gross and Thompson's (2007) model of attention regulation to establish the influence that joy and surprise have on consumers' momentary decisions to engage with ads.

Specification of Lags

Eye-movements, facial expression of emotions and zapping have different response latencies and thus do not occur simultaneously. Coordinated motor actions (zapping) are slower than facial expressions, which are slower than the eyes (Hansen and Hansen 1994). Two questions are therefore when emotions influence IAD and zapping, and whether IAD also influences emotions. We summarize the results of auxiliary analyses to answer the questions and to correctly specify lagged effects in $X^{(1)}$ and $X^{(2)}$ in Equation 4 (Web Appendix B).

We regressed surprise and joy on IAD, and all three of them on zapping, including control variables, for lags of 0 (effect occurs within 250 msec.) to 5 frames (within 1500 msec.). The direction of causality between IAD and emotions is deduced from their correlation and temporal precedence. The results with guidelines for specification of lags were:

1. Levels of surprise and joy affect zapping with a lag of 2 frames (500 to 750 msec.),
2. Velocities of surprise and joy have an instantaneous effect on IAD (0 to 250 msec.),
3. IAD (and AAD) affect zapping with a lag of 2 frames (500 to 750 msec.),
4. IAD has no direct effect on surprise and joy.

These findings are consistent with our theory (Figure 1) and with prior research on the emotion-attention relationship. Finding 2 is in line with evidence that emotionally salient content causes eyes to orient both reflexively and voluntary, between 160 msec. and 320 msec. (1 frame) (Nummenmaa, Hyönä and Calvo 2009). Findings 3 and 4 are in line with theories of saccadic latencies and a separate reaction time experiment (not reported here to save space). Detailing point 4, it states that attention dispersion, by itself, has no direct effect on surprise or joy. This is reasonable because there is no physiological reason why deviations of the eyes from what others are viewing, as captured by IAD, should by themselves induce an emotion within a 250 msec. time window of the data. Thus, theory and empirical evidence jointly point to emotion directing attention in the current context.

Model Estimation

Because of the frailty and hierarchical structure of the model, we estimated it using Bayesian methodology with data augmentation for the Probit model. We used the Forward Filtering Backward Sampling (FFBS) algorithm (Frühwirth-Schnatter 1994) to estimate θ_t, after which ω is straightforward to sample, used separate conditional Bayesian shrinking steps for μ_i and α_c, and a simple Bayesian linear regression step for $\Psi^{(1)}$, $\Psi^{(2)}$, ψ_1, ψ_2 and ψ_3. Elements of the variance-covariance matrix of ε_{ict} were obtained via post-processing. All priors are standard conjugate diffuse priors. The MCMC chain was run for 50,000 iterations on a total of 145,000 observations. The posterior distributions of the parameters of 2,000 draws were extracted, thinning 1 in 10 draws, after a burn-in period of 30,000. Starting values were from the maximum likelihood parameter estimates from independent homogeneous Probit and Log-Linear models. Convergence was checked through Geweke's (1992) z-score, which at -0.18 (p=0.42) did not reject stationarity. As for the individual regression parameters, 60% of the 582 estimates had z-cores within the ±1.96 interval. The other 40% seem stable on a visual check of the trace plots.

RESULTS

Model Diagnostics and Specification

Duration of the ads ranges from 15 to 102 seconds (60 to 408 frames) in our dataset. The number of ads watched by at least one person drops to less than 10 after 120 frames (30 seconds), and to less than 5 after 200 frames. Because this makes model estimation for durations longer than 30 seconds unstable, we truncated the data at 120 frames. This leads to using 81.2% of the total data and provides stable estimates and good convergence of the MCMC iterations from different starting values. Further, the frame-by-frame correlation between joy velocity and absolute velocity, and that between surprise level and velocity was high, above 0.60. This hampered stable estimation. Our solution was to estimate only time-varying intercepts and parameters for the levels of emotions, letting the velocity and absolute velocity parameters have time-invariant effects. Consequently, collinearity is not a major problem in the final model since no pair of included variables had a correlation higher than 0.5. Condition numbers were 3.09 for the zapping model, and 2.87 for the attention model.

Table 3 gives the explained variance of individual, ad, and temporal heterogeneity effects and the sets of explanatory variables, GPR^2 (Gelman and Pardoe 2006). GPR^2 is the equivalent of adjusted R^2 for multilevel models. It requires fitting only the full model, which is convenient given the fact that our proposed model takes more than 300 hours to converge using 120 frames (about four times longer for 240 frames) on a grid server. We estimated the GPR^2 at each level of the hierarchy for the full model in Equation 2. The GPR^2 for the first level of the full model is 63% and confirms that the sets of variables jointly explain a significant portion of the variance in observed zapping. The GPR^2 of 94% for the time-varying component of the model shows the importance of taking temporal heterogeneity of parameters into consideration. The GPR^2 of 19% and 2% for ad and individual heterogeneity, respectively, further shows the importance of the former. In other words, the video ads are fairly different from one another, and viewer demographics have a very low explanatory power in this study (adding interactions between demographics did not improve this). Table 3 shows that ad familiarity is an important predictor of increased avoidance with 18% of relative explained variance. Emotions capture a massive 72% of the relative explanatory power, with all other variables explaining the remaining 10%.

[Insert table 3 about here]

We estimated alternative models with additional emotional measures to assess whether important emotional features or interactions between them were left out, and we tested whether other lag effects that had emerged from the auxiliary analyses would improve the model. This was not the case. Prior to the final analyses, all independent variables were standardized to facilitate comparison of parameter estimates.

Emotional Consequences on Attention and Zapping

Table 4 summarizes the posterior distributions of the parameters. It shows that high levels of joy increase attention concentration (-.032), but that the velocity and absolute velocity of joy are not significant. This supports hypothesis 1: in this situation of high approach motivation, joy prompts concentration of attention rather than exploration. In further support of hypothesis 1, higher levels of surprise also induce attention concentration, in line with its theoretical “halt then reorient” function. Also in support for hypothesis 1, surprise has a stronger influence on attention than joy has (three times more: -.105 versus -.032). The positive velocity effect of surprise on IAD (.006) indicates that fast changes in surprise somewhat attenuate the attention concentration effect due to higher levels of surprise. But, the net effect of surprise level on IAD dominates. These results demonstrate the rapid concentration of attention due to surprise.

[Insert table 4 about here]

In support of hypothesis 2, both emotions also directly reduced zapping, even while controlling for attention concentration (IAD). As predicted by hypothesis 2, the effects of joy on zapping are larger than the effects of surprise (level: -.398 versus -.164, and velocity: -1.818 versus -.168). Interestingly, over and above these effects, joy has an asymmetric effect on zapping: positive changes in joy reduce zapping more than negative changes in joy increase it, with 98% probability. These findings provide strong support our hypotheses.

Figure 2 shows the time-varying intercepts of attention dispersion and zapping. These capture inherent dynamics in the time series, not accounted for by the predictors, but caused by unobserved aspects of the ads. The zapping baseline evolves over time, starting from lower zapping rates in the beginning of the ads. The attention dispersion baseline drops over time from initially higher levels, to rise again after about 20 seconds. As for the dynamic influences of the emotions predicted by emotion regulation theory, joy reduces zapping progressively across exposure: viewers are less likely to zap in the later portion of video ads than in the beginning when experiencing joy. Clearly the first few seconds are critical for ads to captivate viewers' attention: early surprise more strongly reduces zapping than surprise later on. These findings may also partly be due to a selection effect, however, when towards the end the people that do not enjoy the ad have zapped away.

As predicted, attention concentration (lower IAD) indeed reduced zapping. Because the two emotions influence attention (IAD) and zapping, this effect reveals the dual route that emotions take to influence zapping directly and indirectly. Table 3 shows the impact of the emotions on zapping and attention dispersion: 72% of the variation in zapping is explained by the emotion measures (58% by changes in joy alone), while the level of surprise is the biggest predictor of attention concentration, with 29%, out of all 14 variables. The effect of visual complexity-squared (mean centered U-shaped effect) on zapping indicates that both lower and higher levels of visual complexity of the ad increase zapping relative to medium levels, independent of all other factors. The influences of the control variables are consistent with expectations.

[Insert figure 2 about here]

Optimal Emotion-Induced Engagement in Video Ads

Advertisers desire their target customers to view video ads fully and pay concentrated attention. Our model and findings can help assess the extent to which online ads achieve these objectives. Ad development is a highly multifaceted creative process that is difficult to formalize, control and measure quantitatively. Yet, awareness of empirical regularities can support creative ad development. Similar in spirit to the creative templates of quality print ads that focus on rhetorical techniques in ad messages (Goldenberg, Mazursky, and Solomon 1999), we describe trajectories of emotions that increase engagement in video ads. Rather than making specific recommendations about the content of video-ads, our approach more modestly aims to evaluate the intensity and timing of emotions to engage viewers longer and more attentively. The specific rhetorical and other ad message and design techniques to evoke the emotions are outside our scope. Our effort aims to illustrate how the creative process can be supported by knowing the influence that typical sequences of joy and surprise emotions over the course of ads have on concentrating attention and retaining viewers, based on our model and data.

First, we assess how the current set of ads performs to identify 'best-in-class' patterns of emotions. This is done by establishing the average emotional trajectory of each video ad, and plugging it into Equation 4, keeping all other ad attributes fixed. This provides scores for attention dispersion and zapping for that video ad. The inverse of these measures reflect estimated attention <u>concentration</u> on screen and <u>retention</u> of that ad. Figure 3 (top) gives a plot of these two measures for the 28 ads in our dataset. The positive association between attention concentration (vertical axis) and viewer retention (horizontal axis) reflects the impact of the emotions. Four ads (black dots) stand out because of their higher predicted attention concentration and viewer retention. These ads apparently induce effective emotion trajectories. To illustrate, Figure 3 (bottom graph) shows the trajectories of joy for the two best performers: Apple Mac and Bud Light.

[Insert figure 3 about here]

Combining this observed emotion trajectory with the parameter estimates (Table 4) indicates that performance of the Bud Light ad is due to the larger impact of positive changes than negative ones, an increasing trend and a high end-peak of joy. The Apple Mac ad does well for different reasons. It has a higher average level of emotion (37% versus 32%) than Bud Light, and a better ability to concentrate attention via stable and high levels of joy. This reveals how different emotion trajectories can lead to comparably high performance of video ads either directly, explained by the zapping model, or indirectly, evidenced from the IAD model.

The analysis and intuition behind optimal emotion trajectories is provided in Appendix A. Based on this we derive five emotion trajectories, with the first two being "optimal". A *Peak-Valley-Peak*[5] repetition of Joy leads to the highest expected retention of viewers. *Peak-and-Stable* is an optimal non-decreasing trajectory. The Apple Mac ad is an exemplar of the latter trajectory. As a basis of comparison, we identified three other trajectories based on emotion timing studies (Baumgartner, Sujan, and Padgett 1997; Wehrle, Kaiser, Schmidt, and Schrerer 2000; Woltman-Elpers, Wedel, and Pieters 2003). A *Stable-and-Peak* trajectory is typical for certain mystery ads that present the key emotional scene at the finale. It targets classical

[5] To make the benchmarks comparable to the ads, a limit on the peak height was imposed separately for joy and surprise based on the maximum average emotion observed in our dataset: 12% for surprise and 48% for joy.

conditioning of attitudes through emotional reinforcement at the end. A *Linear Increasing* trajectory is based on the idea that people prefer upward sloping emotional sequences more generally. Lastly, since it may be a stretch to assume or desire that ads induce maximal peaks and valleys in emotions at each point in time during a 30 second video ad, we also estimate a milder version of Peak-Valley-Peak. In this template, which we call *Increasing Peak-Valley-Peak,* changes are not as abrupt with a positive trend over time, finishing at the highest level. The Bud Light ad is an exemplar.

We estimate the attention concentration and retention levels of the five emotion trajectories from the parameter estimates and compare them with the observed trajectories of the ads in our data, after normalizing their effects relative to those of the Linear Increasing trajectory (see Figure 4). For example, Bud Light performs about 50% better on retention and about 20% better on attention concentration than the Linear Increasing type. Figure 4 also shows how the four best ads (shown as black dots) compare with the others (white dots) and the five emotion trajectories (grey diamonds). Optimal trajectories are the Peak-and-Stable, best overall in concentrating attention, and the Peak-Valley-Peak pattern, best in retaining viewers.

A novel insight is that the ‘emotional rollercoaster’ caused by the ups and downs of the Peak-Valley-Peak and Increasing Peak-Valley-Peak trajectories reduce attention concentration (relative to Peak-and-Stable), despite ultimately resulting in more viewer retention. Further, since there is no positive effect of decreases in emotion (the asymmetry effect) on attention concentration, attempts to use emotions to increase attention concentration over that of the Peak-Valley-Peak profile will result in lower retention rates. Similarly, any gains in retention above that of the Peak-and-Stable type will result in lower attention concentration. The other three trajectories are dominated on at least one of the two dimensions by the Peak-Valley-Peak and

Peak-and-Stable patterns, with the End-Peak performing worst on both dimensions.

It is noteworthy that most ads in our sample can in theory be improved in terms of attention concentration, retention, or both, by mimicking one of the four prototypical benchmark profiles (but not Stable-and-Peak). Yet, when it comes to improving the time-course of emotions for the four best ads, major gains in viewer retention can only be attained by compromising attention concentration, and vice versa. Our analysis pinpointed this trade-off between attention concentration and viewer retention, at the high end of the attention-retention spectrum (Figure 4). Even though it is challenging to induce the optimal emotion trajectories, there is value in using them as a template in creative ad design. Having identified these emotion trajectories can also inform pretesting practices of video ads. It focuses advertisers and agencies on when, which specific emotions and emotion aspects of video ads need improvement, and what the likely gains of potential improvement are.

[Insert figure 4 about here]

DISCUSSION

Avoidance of television advertising has become a major problem for the advertising industry. This is one reason why advertisers are migrating to the Internet. But, viewers exhibit avoidance behaviors online as well, either via lack of attention concentration or simply by various forms of zapping, clicking, or scrolling the ads. Emotionally engaging video ads are being used to capture and retain target viewers' attention and keep them from zapping. Yet insights into the effectiveness of emotions in attaining these goals, in particular, from moment-to-moment during advertising are limited. To date, few guidelines exist on when to evoke which specific emotions in ads, and no benchmarks exist to evaluate what works well and what not.

We offered the first study, to our knowledge, to examine multiple moment-to-moment emotions via automated facial expression detection and to disentangle the influence of two frequently used emotions, joy and surprise, that are targeted in about two thirds of online ads. The proposed dynamic model yields diagnostic information about the specific moments that trigger consumers to lose concentration and/or to zap. Based on the model, we derive optimal emotion trajectories and compare them to observed trajectories in our data.

With companies like P&G and Unilever leading the way by investing in high-frequency facial emotion tracking of consumers, we believe that our proposed method can be of newfound value to the endeavor of these and other companies, and to the developers of software for automatic emotion recognition from facial expressions, such as eMotion (U. of Amsterdam), used in this research, but also FaceReader (by Noldus) and OKAO (by Omron Corp). This is a promising technology not only for advertising (GfK's -frequency facial emotions measurement lab) but also for the field of human-computer interaction (Sony Cyber Shot line of cameras with "smile shutter" to automatically trigger a snap shot when people smile). The challenge in these applications is to make sense out of the massive amount of data and use it appropriately to predict behavior. Our proposed method provides a first step.

Novel Findings and Implications

We found evidence that, from moment-to-moment during ad exposure, the emotions joy and surprise influence viewer retention directly and indirectly via their influence of attention concentration, which in its turn affects viewer retention. Both joy and surprise led to concentration of attention, which reflects the attention grabbing power of ads at those moments. The attention concentration effects of surprise were much stronger than those of joy. The finding that joy led to attention concentration is counter to predictions based on the broaden-and-build model (Frederickson 1998), but fits with recent findings in psychology. The source of the positive emotion was typically outside (exogenous) of the focal task in prior work on the broaden-and-build model. Gable and Harmon-Jones (2008) however found that (approach-motivated) positive emotions actually focused attention on the source that led to the emotion, i.e., when the positive emotion is "endogenous". We speculate that the attention concentration effect of joy in our study arises because the experience of joy concentrates attention on the source of the emotion, which is in the video ad.

We also found that attention concentration by itself reduced the likelihood of zapping video ads from moment-to-moment, supporting earlier findings (Teixeira, Wedel, and Pieters 2010). It suggests that television ads with the ability to concentrate consumers' visual attention on specific locations in the ad, and thereby reduce heterogeneity in attention, are able to retain consumers effectively. This study uses inter-individual measures of attention concentration at a quarter-second rate. While these are easy to compute and effective, analog measures of intra-individual attention concentration that assess the extent to which the same individual focuses on a stimulus over time would be of interest to develop and test in future research.

We found too that surprise improved attention concentration more than joy did, and joy improved viewer retention more than surprise did, revealing the dual routes to ad effectiveness that these two related but distinct emotions play. These findings provide guidelines for advertisers interested in grabbing and retaining their target customers' attention at specific moments during an ad, and in aiming to retain their customers.

Furthermore, the velocity of both emotions influenced viewer retention; for joy the effect was even larger than that of its level. This is evidence of the importance of momentary changes in emotions on attention and decisions, which seem not to have been previously documented. We see our study as a first step, and hope that future work will further establish a firm empirical and theoretical basis for these effects.

The emotion trajectories that this study identified may serve as guides in ad development. A Peak-and-Stable trajectory appears particularly useful for ads seeking to maximize attention concentration; a Peak-Valley-Peak trajectory when maximizing viewer retention is the goal. Improving ads in the market place may eventually be a trade-off between further increases of attention concentration or of viewer retention. Our prototypical emotional profiles are useful in particular for ads in which editing and scene permutation in the stages of creative design can effectively be used to change the location of emotional scenes without compromising on other important holistic attributes (e.g. narrative, aesthetics, persuasive argumentation). One such case is movie trailer ads[6] in which multiple short snippets of the movie is put together to generate attention. Our approach can be of newfound value in such and potentially other cases.

TiVo President Tom Rogers emphasized that: "We are already processing a billion pieces of second-by-second data a day that demonstrates exactly what commercials are seen and which are not."[7] The current study demonstrates that our model calibrated on zapping data at even higher temporal frequency supplemented with facial expression and attention data provides the required insights to help improve the predictions of which video ads consumers will see, or not.

APPENDIX A

To assess the optimal time-course of emotions, we break down the emotion trajectories into their units, namely peak (increase), valley (decrease) or stable (flat) moments. We focus on joy for simplicity of exposure. We discretize exposure time and emotional intensity to one unit.

[6] Movie trailers are popular on the Internet. In 2010 estimates of 10 billion videos watched online, movie trailers rank third, after news and user-generated (http://en.wikipedia.org/wiki/Trailer_(film), accessed February 3, 2011.
[7] http://www.huffingtonpost.com/jack-myers/tv-industry-faces-ad-avoi_b_136421.html, accessed February 22, 2011.

We can now calculate all possible trajectories from the origin (0) to the end (T=120), where we assume the emotion is absent at t=0. From Figure A.1, there are two possible trajectories: an upward increase, denoted by p (for peak), or a flat trajectory denoted by s (for stable). Using the model parameters the estimated utility from each option is calculated below. Let

$${}^{\text{change}}\theta_t = {}^{\text{velocity}}\theta_t + {}^{\text{absolute velocity}}\theta_t \,,$$

utility$\{p\}$ = ${}^{\text{level}}\theta_T * 1 + {}^{\text{change}}\theta_T * 1 = \mathbf{3.711}$

utility$\{s\}$ = 0

Clearly, the peak trajectory is preferred. This analysis eliminates a null emotional trajectory as a potential optimal solution.

We now compare trajectories for a case of two discrete times and two emotional levels. Again, such a situation may occur in practice if the creative process has yielded two joy-evoking events that need to be placed in the video ad. As before, a null trajectory need not be considered, as well as any trajectory that doesn't end at the highest terminal level (T,1), since this is the last opportunity in the ad to 'collect' utility from higher joy levels. This leaves us with three potential trajectories combining 'peak' or 'stable' segments, as can be seen in Figure A.1.B. Here the assumption is that the emotion is absent at t=0 and it is possible to reach maximal emotion level in half the ad time. As before, using the parameter estimates of the retention model, we calculate the estimated contributions to utility of each trajectory as follows:

utility$\{p\text{-}s\}$ = ${}^{\text{level}}\theta_{T/2} * 1 + {}^{\text{change}}\theta_{T/2} * 1 + {}^{\text{level}}\theta_T * 1 = \mathbf{4.118}$

utility$\{p\text{-}s\}$ = $({}^{\text{level}}\theta_{T/2} + {}^{\text{change}}\theta_{T/2}) * 1/2 + {}^{\text{level}}\theta_T * 1 + {}^{\text{change}}\theta_T * 1/2 = 3.915$

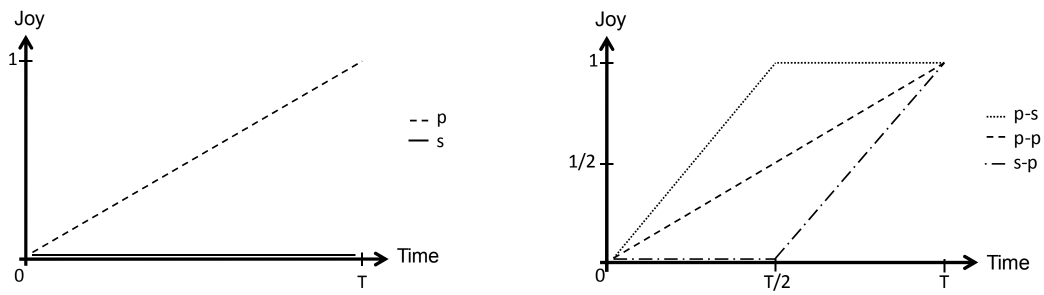
utility$\{s\text{-}p\}$ = ${}^{\text{level}}\theta_T * 1 + {}^{\text{change}}\theta_T * 1 = 3.711$

Since change parameters don't vary over time, after canceling terms, it is evident that the trajectory with highest predicted retention, p-s, benefits uniquely from a higher parameter for the level of joy at T/2. Discretizing the space of possible trajectories to three time periods and three emotion levels, an analogous analysis permits 10 paths with only up or flat segments, starting at (0,0) and ending at (T,1).

FIGURE A.1

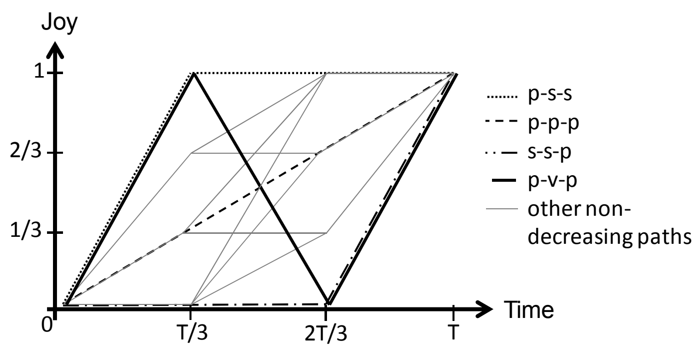
SPACE OF FEASIBLE EMOTIONAL TRAJECTORIES

A) One time/one level B) Two time/two levels



Note: s=stable, p=peak, v=valley

C) Three time/three levels



Examples are provided in Figure A.1.C. Again, these may have practical relevance if the creative process has identified three emotional events that need to be placed in the ad. Now a series of trajectories with decreases (*v* for valley) in emotion levels becomes feasible, all of which are contained with the convex hull of the *p-s-s* and *s-s-p* profiles. Utilities of all trajectories are calculated, the most important ones being:

utility{*p-s-s*} = ${}^{level}\theta_{T/3} * 1 + {}^{change}\theta_{T/3} * 1 + {}^{level}\theta_{2T/3} * 1 + {}^{level}\theta_{T} * 1 = 4.516$

utility{*p-p-p*} = $({}^{level}\theta_{T/3} + {}^{change}\theta_{T/3}) * 1/3 + {}^{level}\theta_{2T/3} * 2/3 + {}^{change}\theta_{2T/3} * 1/3 + {}^{level}\theta_{T} * 1 + {}^{change}\theta_{T} * 1/3 = 4.115$

utility$\{s\text{-}s\text{-}p\}$ = ${}^{\text{level}}\theta_T * 1 + {}^{\text{change}}\theta_T * 1 = 3.711$

utility$\{p\text{-}v\text{-}p\}$ = ${}^{\text{level}}\theta_{T/3} * 1 + {}^{\text{change}}\theta_{T/3} * 1 + ({}^{\text{velocity}}\theta_{2T/3} - {}^{\text{abs. velocity}}\theta_{2T/3}) * -1 + {}^{\text{level}}\theta_T * 1 + {}^{\text{change}}\theta_T * 1 = \mathbf{7.012}$

The non-decreasing path with the highest predicted viewer retention is, similar to the previous cases, the trajectory with the quickest increase in joy and subsequent stable delivery until the end (*p*-*s*-*s*). But the optimal trajectory has the fastest increase of joy, followed by the fastest decrease to zero and then increasing again (*p*-*v*-*p* in Figure A.1.C). Comparing all alternatives can show this. The path *p*-*s*-*s* leads to less retention than the optimal *p*-*v*-*p* path as long as the inequality ${}^{\text{abs. velocity}}\theta > {}^{\text{level}}\theta_{2T/3}/2$ holds. In essence, the benefit comes from the asymmetry of the change in joy, which induces higher utility through positive changes than that due to subsequent decrease in the emotion. But, of course practical and executional considerations may cause the creative to adopt the non-decreasing emotional path.

When the number of discrete times and emotion levels are increased, *p*-*s*-*s* and *p*-*v*-*p* type trajectories are still the non-decreasing, respectively, unconstrained optimal trajectories, because of the Markovian property of the model. Generalizing this multiple emotional levels and multiple time periods, these trajectories correspond to the vectors (0,1,1,1…) and (0,1,0,1,…), respectively. The former corresponds to a pure maximal delivery of emotion via an initial peak and stable delivery at the maximum level. The latter corresponds to the path with highest variation in the emotion delivery with an up-down-up repetition. It is an example of the peak-and-end rule for the relationship between moment-to-moment affective evaluations and retrospective evaluation of an episode. There, the peak-and-end rule applies to evaluations of the past being disproportionally influenced by two singular moments. In our context of self-exposure, the evidence shows the importance of providing emotional changes that momentarily increase viewing retention and, under the reasonable assumption that viewers might be integrating past moments, this is consistent with a high evaluation due to the peak-and-end rule.

REFERENCES

Bartlett, Marian Stewart, Joseph C. Hager, Paul Ekman and Terrence J. Sejnowski (1999), "Measuring Facial Expressions by Computer Image Analysis," *Psychophysiology*, 36, 253-63.

Baumgartner, Hans, Mita Sujan, and Dan Padgett (1997), "Patterns of Affective Reactions to Advertisements: The Integration of MTM Responses into Overall Judgments," *Journal of Marketing Research*, 34 (May), 219-32.

Calvo, Manuel G. and Peter J. Lang (2004), "Gaze Patterns When Looking at Emotional Pictures: Motivationally Biased Attention," *Motivation and Emotion,* 28 (3), 221–43.

Clarkson, Douglas B., Chris Fraley, Charles Gu, and James O. Ramsay (2005), *S+Functional Data Analysis User's Guide.* New York: Springer.

Cohen, Ira, Nicu Sebe, Ashutosh Garg, Lawrence S. Chen, and Thomas S. Huang (2003), "Facial Expression Recognition from Video Sequences: Temporal and Static Modeling," *Computer Vision and Image Understanding*, 91 (1-2), 160-87.

Darwin, Charles (1872), *The Expression of the Emotions in Man and Animals*, London: Murray.

Derbaix, Christian 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 (November), 470-9.

Ekman, Paul and Wallace V. Friesen (1978), *The Facial Action Coding System.* Palo Alto, CA: Consulting Psychologists Press.

---, (1999), "Facial expressions," in *Handbook of Cognition and Emotion*, Tim Dalgleish and Michael J. Power, eds. Chichester: John Wiley and Sons Ltd., 301-20.

Elkin, Tobi (2010), "Consumer Packaged Goods Sector Taps into Online Video," (accessed February 9, 2010), [available at http://www.emarketer.com].

Fredrickson, Barbara L. (1998), “What Good Are Positive Emotions?” *Review of General Psychology*, 2 (3), 300–19.

--- and Christine Branigan (2005), “Positive Emotions Broaden the Scope of Attention and Thought–action Repertoires,” *Cognition and Emotion*, 19 (3), 313–32.

Frühwirth-Schnatter, Silvia (1994), “Data Augmentation and Dynamic Linear Models,” *Journal of Time Series Analysis*, 15(2), 183-202.

Gable, Philip A. and Eddie Harmon-Jones (2008), "Approach-Motivated Positive Affect Reduces Breadth of Attention," *Psychological Science*, 19 (5), 476-482.

Gelman, Andrew and Iain Pardoe (2006), “Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models,” *Technometrics,* 48(2), 241-51.

Geweke, John (1992), "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments." in Bayesian Statistics 4, J. M. Bernardo, J. Berger, A. P. Dawid and A. F. M. Smith, eds. Oxford, UK: Oxford Univ. Press, 169-93.

Goldenberg, Jacob, David Mazursky, and Sorin Solomon (1999), “The Fundamental Templates of Quality Ads,” *Marketing Science*, 18 (3), 333-51.

Griskevicius, Vladas, Michelle Shiota, and Stephen Nowlis (2010), “The Many Shades of Rose-Colored Glasses: An Evolutionary Approach to the Influence of Different Positive Emotions.” *Journal of Consumer Research*, 37 (August), 238-50.

Gross, James J. and Ross A. Thomson (2007), "Emotion Regulation, Conceptual Foundations," in: *Handbook of Emotion Regulation*, James J. Gross (Ed.). Guilford Publications.

Grover, Ronald and Joe Fine (2006), “The Sound of Many Hands Zapping,” (accessed June 10, 2010), [available at http://www.businessweek.com/magazine/content/06_21/b3985063.htm].

Gustafson, Paul. and S. Siddarth (2007), “Describing the Dynamics of Attention to TV

Commercials: A Hierarchical Bayes Analysis of the Time to Zap an Ad," *Journal of Applied Statistics,* 34(5), 585-609.

Hansen, Christine H. and Ranald D. Hansen (1994), "Automatic Emotion: Attention and Facial Efference," in *The Heart's Eye: Emotional Influences in Perception and Attention*, P.M. Niedenthal and S. Kitayama, eds. San Diego: Academic Press, 217-43.

Hajcak Greg and Doreen M. Olvet (2008), "The Persistence of Attention to Emotion: Brain Potentials During and After Picture Presentation," *Emotion*, 8(2), 250–5.

Hsee, Christopher K. and Robert P. Abelson (1991), "Velocity Relation: Satisfaction as a Function of the First Derivative of Outcome over Time," *Journal of Personality and Social Psychology*, 60 (March), 341-47.

Kahneman, Daniel (2000), "*Evaluation by Moments: Past and Future*," in Choices, Values, and Frames, D. Kahneman and A. Tversky, eds. Cambridge, UK: Cambridge Univ. Press, 673–92.

Lemmink, Jos and Jan Mattsson (1998), "Warmth During Non-productive Retail Encounters: the Hidden Side of Productivity," *International Journal of Research in Marketing*, 15, 505-17.

Meyer, Wulf-Uwe, Rainer Reisenzein, and Achim Schützwohl (1997), "Toward a Process Analysis of Emotions: The Case of Surprise," *Motivation and Emotion*, 21 (3), 251-74.

Mihaylova, Milena, Vladislav Stomonyakov and Angel Vassilev (1999), "Peripheral and Central Delay in Processing High Spatial Frequencies: Reaction Time and VEP Latency Studies," *Vision Research, 39,* 699-705.

Nummenmaa, Lauri, Jukka Hyönä, and Manuel G. Calvo (2009), "Emotional Scene Content Drives the Saccade Generation System Reflexively," *Journal of Experimental Psychology: Human Perception and Performance,* 35(2), 305-23.

Olney, Thomas J., Morris B. Holbrook, and Rajeev Batra (1991), "Consumer Responses to

Advertising: The Effects of Ad Content, Emotions, and Attitude Toward the Ad on Viewing Time," *Journal of Consumer Research*, 17 (March), 440-53.

Olsen, G. Douglas and John W. Pracejus (2004), "Integration of Positive and Negative Affective Stimuli," *Journal of Consumer Psychology*, 14 (4), 374-84.

Plutchik, Robert (1980), *Emotion: A Psychoevolutionary Synthesis*. New York: Harper & Row Pub.

Ramsay, James O., Giles Hooker, and Spencer Graves (2009), *Functional Data Analysis with R and MATLAB*. New York: Springer

Sueyoshi, Glen. T. (1995), "A Class of Binary Response Models for Grouped Duration Data," *Journal of Applied Econometrics*, 10(4), 411-31.

Teixeira, Thales, Michel Wedel, and Rik Pieters (2010), "Moment-to-moment Optimal Branding in TV Commercials: Preventing Avoidance by Pulsing," *Marking Science*, 29(5), 783-804.

Tibshirani, Robert, Guenther Walther and Trevor Hastie (2001), "Estimating the Number of Clusters in a Dataset via the Gap Statistic," *Journal of the Royal Statistics Society B*, 411-23.

Vakratsas, Demetrios and Tim Ambler (1999), "How advertising works: What do we really know?" *Journal of Marketing*, 63(1) 26–43.

Wehrle, Thomas, Susanne Kaiser, Susanne Schmidt, and Klaus R. Scherer (2000), "Studying the Dynamics of Emotional Expression Using Synthesized Facial Muscle Movements," *Journal of Personality and Social Psychology*, 78(1), 105-19.

Woltman-Elpers, Josephine L. C. M., Michel Wedel, and Rik G. M. Pieters (2003), "Why Do Consumers Stop Watching TV Commercials? Two Experiments on the Influence of Moment-to-Moment Entertainment and Information Value," *Journal of Marketing Research*, 40 (November), 437–53.

Yiend, Jenny (2010) "The Effects of Emotion on Attention: A Review of Attentional Processing of Emotional Information," *Cognition and Emotion*, 24 (1), 3-47.

TABLE 1

ADS AND EMOTIONS

Brand	Title or Tag line	Product category	Average Intensity: Joy	Average Intensity: Surprise	Average Positive Intensity: Surprise	Participants who felt emotion for more than 1 second: Joy	Participants who felt emotion for more than 1 second: Surprise
Bud Light	Swear jar	Alcoholic Beverage	32%	5%	23%	24	10
Miller	Genuine draft: There's a party in there.	Alcoholic Beverage	8%	5%	18%	8	7
K-fee	The drive	Beverage	14%	4%	15%	13	3
Rockstar	Mayhem festival	Beverage	11%	5%	19%	6	4
Therabreath	Cigarette breath?	Hygiene	15%	4%	24%	11	6
Dentyl	New Dentyl PH Exhilaration mouthwash	Hygiene	8%	3%	18%	11	2
Apple Mac	featuring Mr. Bean	Computer	37%	6%	23%	30	8
Dell	Tell us what you want.	Computer	17%	2%	15%	13	4
Nivea	Body lotion for men	Beauty	14%	4%	15%	9	3
Dove	Real Beauty / Self-esteem fund	Beauty	3%	6%	19%	6	6
Quit	Quitline	NGO	11%	3%	21%	5	4
Gift of Life and Breath	Run for the gift of life and breath.	NGO	4%	5%	16%	4	4
Whitecastle	True castle stories.	Fast Food	26%	3%	18%	22	3
Carls Jr	The all new…	Fast Food	20%	1%	12%	9	3
Sony Playstation 2	Different places. Different rules.	Videogame	12%	2%	10%	12	2
Xbox	360 Elite	Videogame	8%	4%	20%	6	4
Vodaphone	Mayfly. Make the most of now.	Telecom	10%	4%	22%	10	8
British Telecom	Come back	Telecom	9%	2%	14%	8	3
Heinz	Takes a while to come out.	Condiments	12%	6%	20%	7	8
Hunts	Yeah. It's that good.	Condiments	13%	3%	14%	12	3
Mountain Dew	Do the Diet Dew.	Beverage	12%	5%	20%	12	9
Sunkist	The original orange.	Beverage	11%	8%	21%	10	7
K-fee	The beach	Beverage	18%	6%	22%	17	6
Rockstar	Party like a rock star.	Beverage	12%	6%	14%	10	5
Mercator	Because you're likely to move out some day.	Financial services	29%	3%	18%	20	5
Lincoln Insurance	Lincoln beats the competition.	Financial services	9%	3%	22%	3	2
Tide	It's gotta be Tide.	Laundry detergent	11%	7%	19%	13	9
Clorox	Magic of new Clorox 2	Laundry detergent	18%	2%	21%	13	4

Note- Average Intensity is the classification accuracy [0-100%] of the facial expression while Average Positive Intensity is only calculated for non-zero instances.

TABLE 2

DATA DESCRIPTION

Variable	Variation units	Mean	SD	Minimum	Maximum
Zapping	across 28 ads	48%	21%	10%	86%
Zapping	across 50 ind.	48%	21%	11%	93%
Emotion:					
Joy level	ad, time, ind.	15.4%	32.0%	0%	100.0%
Joy velocity	ad, time, ind.	0%	2.0%	-14.6%	14.6%
(Joy) Absolute velocity	ad, time, ind.	0%	1.8%	0%	14.6%
Surprise level	ad, time, ind.	4.4%	19.1%	0%	100.0%
Surprise velocity	ad, time, ind.	0%	1.4%	-14.5%	14.1%
Attention dispersion:					
Individual dispersion (pixels^2)	ad, time, ind.	3481	1764	0	91204
Aggregate dispersion (pixels^2)	ad, time, ind.	6837	3276	696	24860
Aggregate×Individual dispersion	ad, time, ind.	4.6×10^5	5.0×10^5	0	5.3×10^6
Control variables:					
Visual complexity^2 (Kbytes^2)	ad, time	2.0×10^4	2.9×10^4	3.6×10^1	3.7×10^5
Brand presence (present = 1)	ad, time	23.6%	42.5%	0	1
Brand duration (seconds)	ad, time	4.6	9.5	0	87
Brand cardinality (order: 1,2,…)	ad, time	1.2	3.4	0	39
Participant age (years)	ind.	22	4.7	18	50
Participant gender (male = 1)	ind.	53.4%	.50	0	1
Ad length (seconds)	ad	43	18.5	15	100(60)
Ad familiarity (familiar = 1)	ad, ind.	7.6%	.26	0	1
Brand familiarity (familiar = 1)	ad, ind.	59.0%	.49	0	1

Note- Zapping statistics are for the complete, non-truncated data irrespective of viewing termination time. Emotion measures are summary statistics of the intensity of the facial expression on a 0 to 100% scale across ad, time and individuals.

TABLE 3

MODEL COMPARISON AND IMPORTANCE
OF VARIABLES IN ZAPPING MODEL

	Hierarchy			
	Utility	Individual	Ad	Temporal
Gelman and Pardoe R^2	62.7%	2.2%	18.9%	93.8%
	Relative importance of variables (%)			
Ad familiarity	18			
Emotions	62			10
Attention	1			
Ad features			4	
Demographics		2		
Other covariates	3			

Note- Covariates: Video Brand, Audio Brand, Size, Duration, Cardinality, SCSQ. In time-varying levels model with all three sources of heterogeneity

TABLE 4

EFFECTS OF JOY AND SURPRISE ON ATTENTION AND ZAPPING

Parameter	Zapping Decision					Attention Dispersion (IAD)				
	Mean	SE	Percentiles of the Posterior Distribution			Mean	SE	Percentiles of the Posterior Distribution		
			5%	50%	95%			5%	50%	95%
Intercept†	**-3.193**	**.147**	-3.440	-3.193	-2.959	-.004	.009	-.018	-.004	.010
Emotion:										
Joy level†	**-.398**	**.056**	-.494	-.397	-.308	**-.032**	**.002**	-.035	-.032	-.028
Joy velocity	**-1.818**	**.142**	-2.006	-1.852	-1.541	.000	.001	-.003	.000	.002
(Joy) Absolute velocity	**-1.418**	**.130**	-1.591	-1.452	-1.156	-.002	.002	-.004	-.002	.000
Surprise level†	**-.164**	**.046**	-.239	-.164	-.090	**-.105**	**.005**	-.112	-.106	-.096
Surprise velocity	**-.168**	**.017**	-.196	-.167	-.141	**.006**	**.001**	.003	.006	.008
Attention dispersion:										
Individual dispersion (IAD)	**.050**	**.017**	.024	.050	.079					
Aggregate dispersion	.023	.017	-.005	.023	.050					
Aggreg.*Indiv. dispersion	-.015	.011	-.032	-.015	.004					
Control variables:										
Visual complexity^2	**.033**	**.013**	.012	.033	.054	**-.017**	**.002**	-.020	-.017	-.014
Brand presence (p = 1)	**.033**	**.016**	.007	.033	.058	**.020**	**.002**	.017	.020	.024
Brand duration (sec.)	**.071**	**.024**	.032	.072	.111	**.027**	.003	.022	.027	.031
Brand cardinality	-.025	.018	-.054	-.025	.004	.003	.002	-.001	.003	.006
Participant age	-.092	.058	-.185	-.094	.004	**-.034**	**.002**	-.038	-.034	-.031
Participant gender (m = 1)	.016	.057	-.074	.016	.105	**.064**	**.003**	.058	.064	.069
Ad length	**-.109**	**.059**	-.210	-.106	-.018	**-.011**	**.002**	-.014	-.011	-.008
Ad familiarity (f = 1)	**1.000**	**.293**	.524	.998	1.496	**.031**	**.002**	.028	.031	.035
Brand familiarity (f = 1)	**-.104**	**.056**	-.194	-.104	-.010	**-.009**	**.002**	-.011	-.009	-.006
Variance of error term	1	0				.211	.018			
Covariance of error term	.0001	.002				.0001	.002			

Note- Bolded estimates have one-sided 95% posterior confidence intervals that do not contain zero; † indicates parameters averaged over time for Zapping model. To reduce skewness of IAD (a squared Euclidian distance), the Euclidian distance was used.

FIGURE 1
GROSS AND THOMSON (2007) MODEL OF EMOTION REGULATION

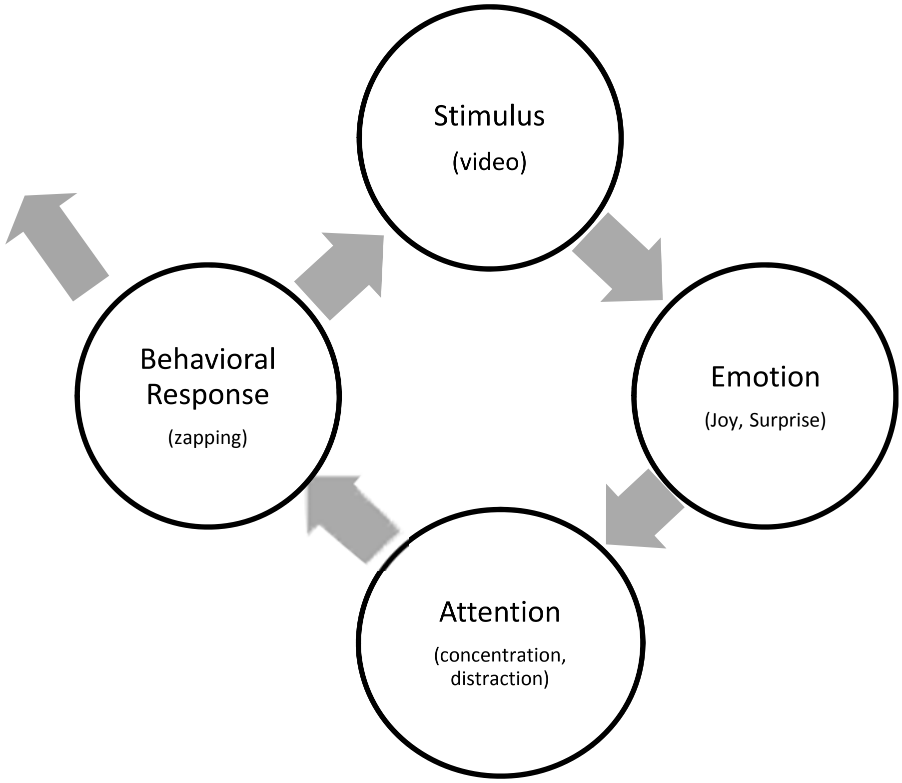
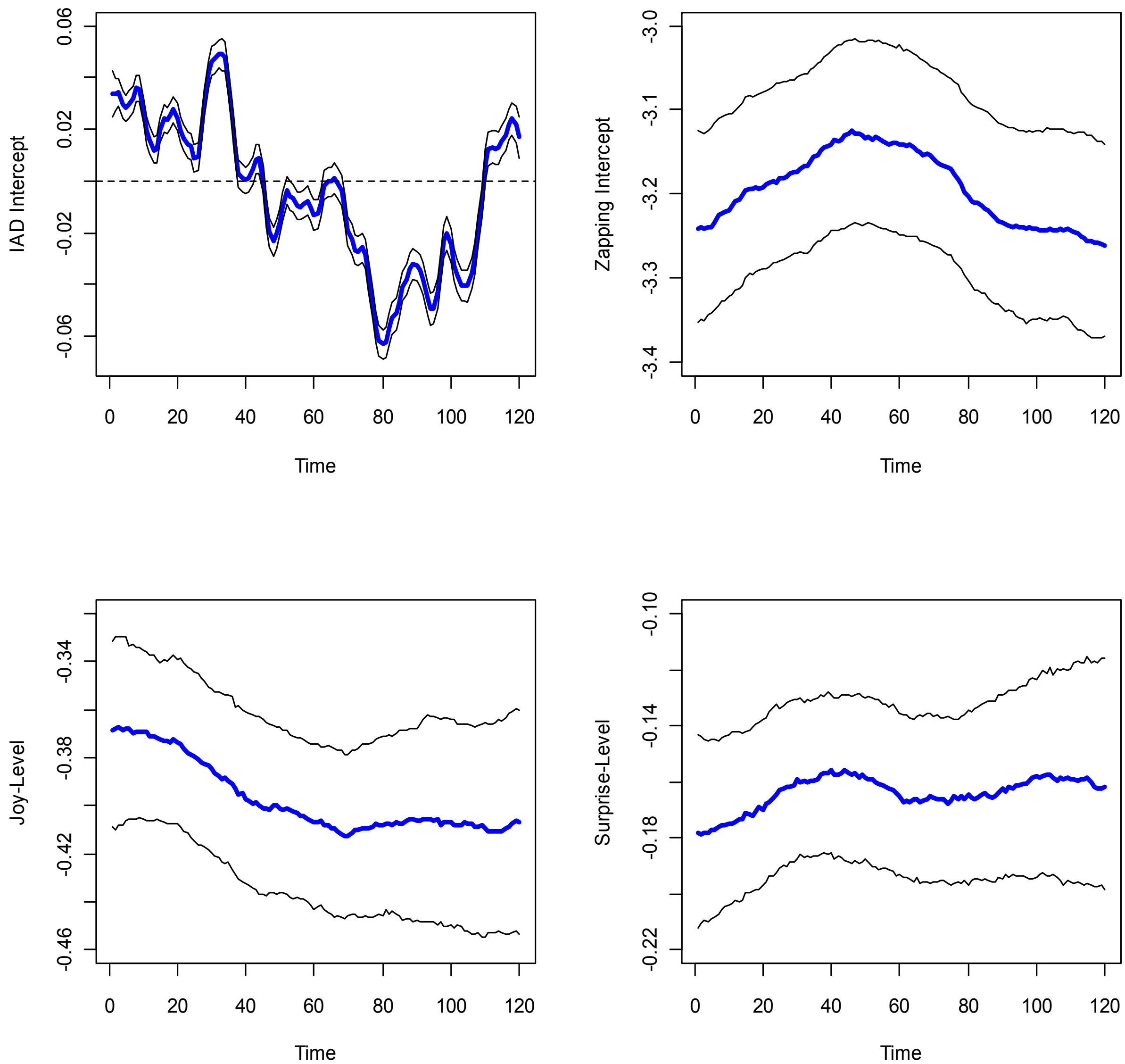


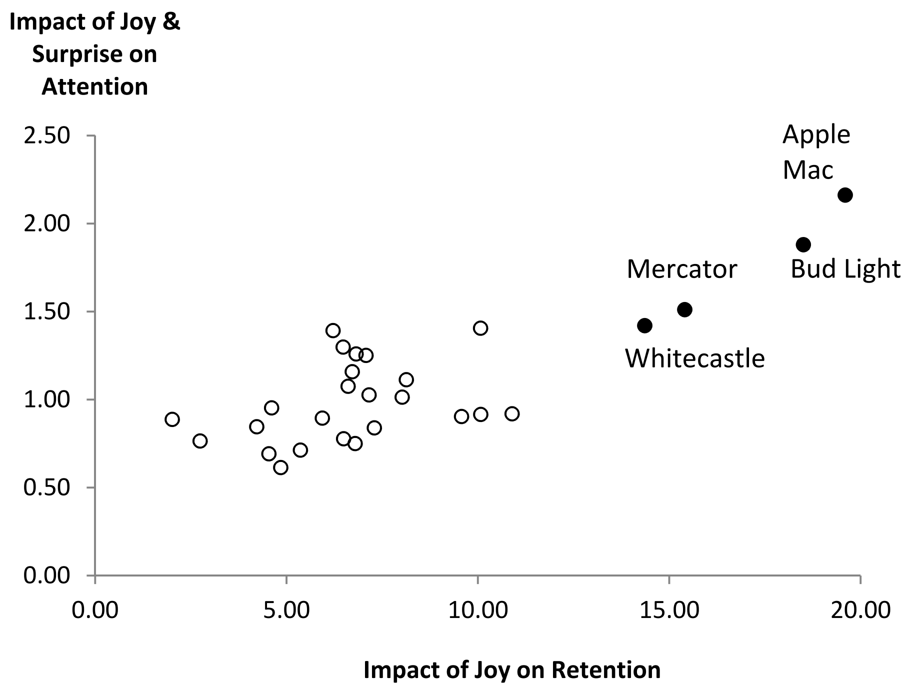
FIGURE 2
MOMENT-TO-MOMENT BASELINE ATTENTION DISPERSION AND ZAPPING RATE (TOP) AND EMOTION EFFECTS ON ZAPPING (BOTTOM)

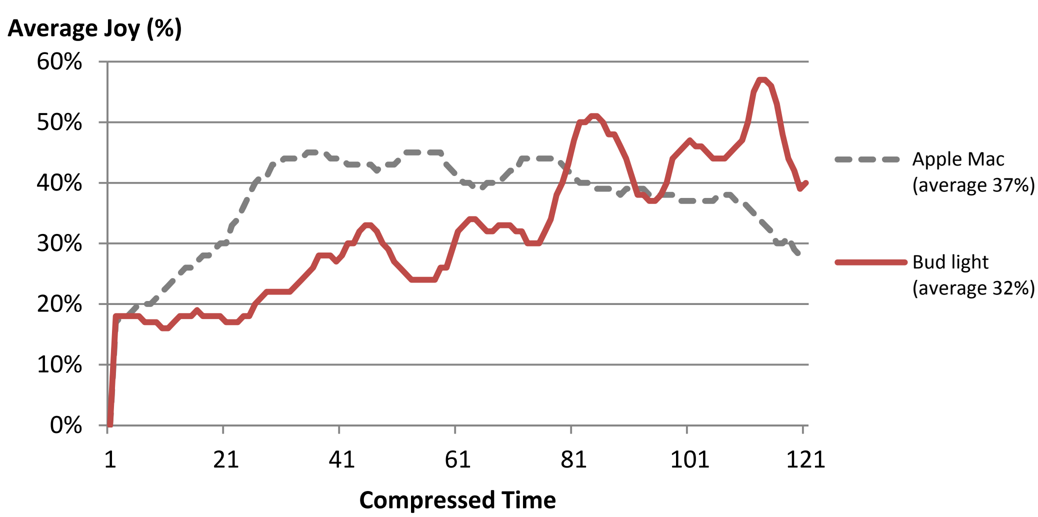


Note- Individual Attention Dispersion is standardized to zero mean.
Thick line is posterior median and thin lines are 1st and 3rd posterior quartiles.

FIGURE 3

IMPACT OF EMOTIONS ON ATTENTION AND RETENTION FOR ALL 28 ADS (TOP) AND AVERAGE EMOTIONAL PROFILE OF JOY FOR TWO BEST ADS (BOTTOM)

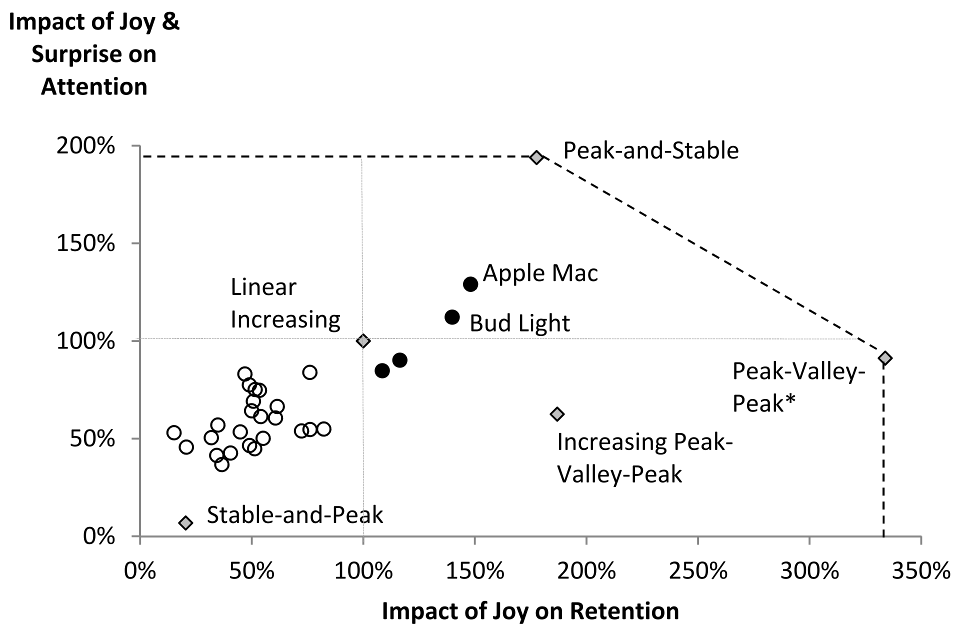




Note- Apple Mac ad is featuring Mr. Bean and Bud Light is intitled Swear Jar.

FIGURE 4

PERFORMANCE OF THE ADS RELATIVE TO EMOTION TRAJECTORIES



Note- Impact on retention of joy for Peak-Valley-Peak is divided by two to fit the graph.

WEB APPENDIX A

CLUSTERING TO FIND EMOTION FEATURES

Here we details the steps to using functional cluster analysis followed by an algorithm to detect the optimal number of clusters, in order to find a small set of variables that capture the most discriminate and pervasive emotional patterns of joy and surprise exhibited in our data. The cluster analysis operates on the functional curves produced by FDA (Ramsay Hooker and Graves 2009).

We first separate the functional response curves by commercial, and cluster them using a hierarchical Ward procedure. The Ward method minimizes the sum of within-cluster variance, which is consistent with the FDA fitting procedure. The distance metric used is defined to be the integral of squared Euclidian distances between the functional curves $e_i(t)$ for each pair of viewers, in the domain represented by the start of the ad and the earliest zapping time between the two. The idea is to compare emotional responses only during the time in which both viewers actually saw the commercial and could have emotions. Since the domain of integration varies by viewer pairs, we standardize the integral to the domain's length. Also, to avoid that at a certain period a large difference between two viewers (emotional spike) has a disproportional effect on the distance measure compared to no emotion, we apply the double log transform ($\ln\{-\ln\{ e_i(t)\}\}$) to the functional values, effectively compressing the values away from zero. If we were to cluster the data solely based on this distance metric of level curves, we could potentially lose diagnostic information regarding derivatives. So, we apply the same distance measure described above to the fitted 1st derivatives produced by FDA of the data. We stopped at 1st derivatives because 2nd derivatives did not add value. In the clustering algorithm, we use a weighted average (convex combination) of the functional levels as well as the functional derivatives, with weights inversely proportional to the range of each distance. This keeps the relative importance of derivatives and levels approximately equal. As recommended by Ramsay et. al (2009), the functions and not the (discrete) data vectors were clustered.

The clustering across viewers is done separately by commercial and emotion, and the number of clusters is defined by means of the Gap Statistic (Tibshirani et al. 2001). This method was chosen because (1) it is based on the same measure of fit as the Ward's method (Sum of Squared Distances, SSD), (2) it uses a prior-type (null reference) distribution, (3) contrary to many other optimal number of clusters methods, it can detect no clustering (or 1 cluster), and (4) the method is fairly robust and performs better than many other techniques (see Tibshirani et al. 2001). Essentially, the Gap statistic compares the log of SSD as a function of the number of clusters with the expected measure under a null distribution and chooses the number of clusters in which the former decays faster than the latter, for the first time. Agnostically choosing the uniform distribution for the location of the functional level and derivatives values, the Ward+Gap procedure generates between two and five clusters for the joy ads, between two and three clusters for the surprise ads and mostly one cluster for the neutral ads. Differences in emotional responses (level, derivative, patterns) aid in choosing measures that can discriminate amongst ads. Only level and velocity were incorporated in the final model as changes in the second derivatives (acceleration) showed no differences between clusters.

References

Ramsay, James O., Giles Hooker, and Spencer Graves (2009), *Functional Data Analysis with R and MATLAB*. New York: Springer

Tibshirani, Robert, Guenther Walther and Trevor Hastie (2001), "Estimating the Number of Clusters in a Dataset via the Gap Statistic," *Journal of the Royal Statistics Society B*, 411-23.

WEB APPENDIX B

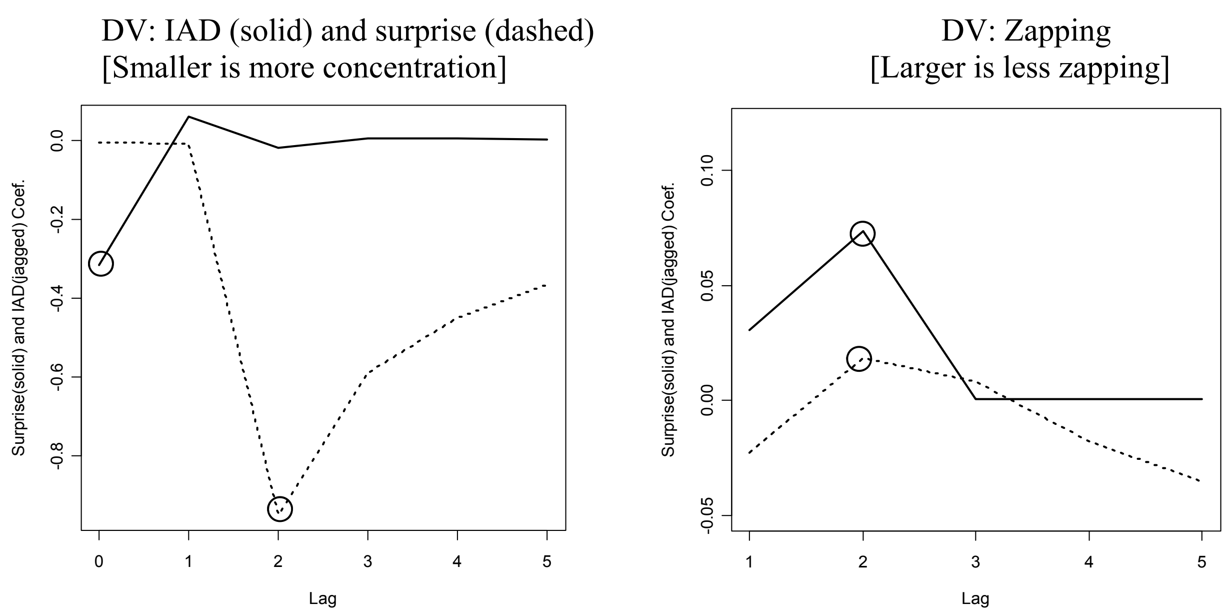
ASSESSING DIRECTION OF CAUSALITY AND LATENCY RESPONSE

This appendix describes the procedure to estimate mean latencies of emotions impact on IAD and on zapping. First, we assess the direction of causality between IAD and emotions. That is, whether emotions drive, as we assume with current emotion literature, attention, vice-versa or both. Assume that at time $t = T_1$ the video ad is shown and has an effect on IAD (and on AAD). Because facial expressions are much slower than the eyes (Hansen and Hansen 1994), if video is also to cause an emotional reaction, it will cause a reaction at time $t = T_1+t'$. The direct effect on the decision to zap the video should be even slower, but is unknown. Now, allowing for the possibility of these effects on surprise/joy and IAD to propagate, it becomes an empirical question to measure the appropriate latency at which this can be picked up as a statistical correlation.

To find the appropriate latency, we regress each variable, surprise/joy or IAD, on the other, controlling for other conjectured commercial, individual and brand-specific effects, for various lags, from 0 (effect occurs within 250 msec.) to lag 5 (within 1500 msec.). All latency findings hold for both emotions but are only shown for surprise, for parsimony.

The lag with the highest regression parameter best explains the influence of the regressor on the regressand and indicates the appropriate average peak latency response. Figure B.1 (left) presents the parameter effects of surprise on IAD (solid line) and of IAD on surprise (dashed line) for simultaneous and lag 1 to 5 models. All values are significant at 5% and similar for joy. Notice that the best fitting model indicates that surprise is related to changes in IAD within a 250 msec. (lag 0) while IAD is more strongly related to changes in surprise after a lag of 2 frames (500 to 750 msec.). Given this longer latency of IAD on surprise, it is plausible that the stimulus caused changes in IAD and this propagated to create an emotional feedback effect from surprise to IAD, and back to surprise. On the other hand, controlling for the direct stimulus effect on surprise, it would seem very unlikely that the mere deviation of a viewer's eyes from the rest of the viewers (higher IAD) could systematically activate an emotion strong enough to evoke a facial expression. These findings are in line with Nummenmaa, Hyönä and Calvo (2009) who report evidence that during complex scene perception (as we have here), emotional content causes the eyes to orient (1st saccade) both reflexively and voluntarily starting at 160 msec. and peaking at 320 msec. Given the lack of any evidence for a fast enough causal effect, we abandoned the conjecture of a direct effect of IAD on surprise or joy and only propose to model the reverse effect.

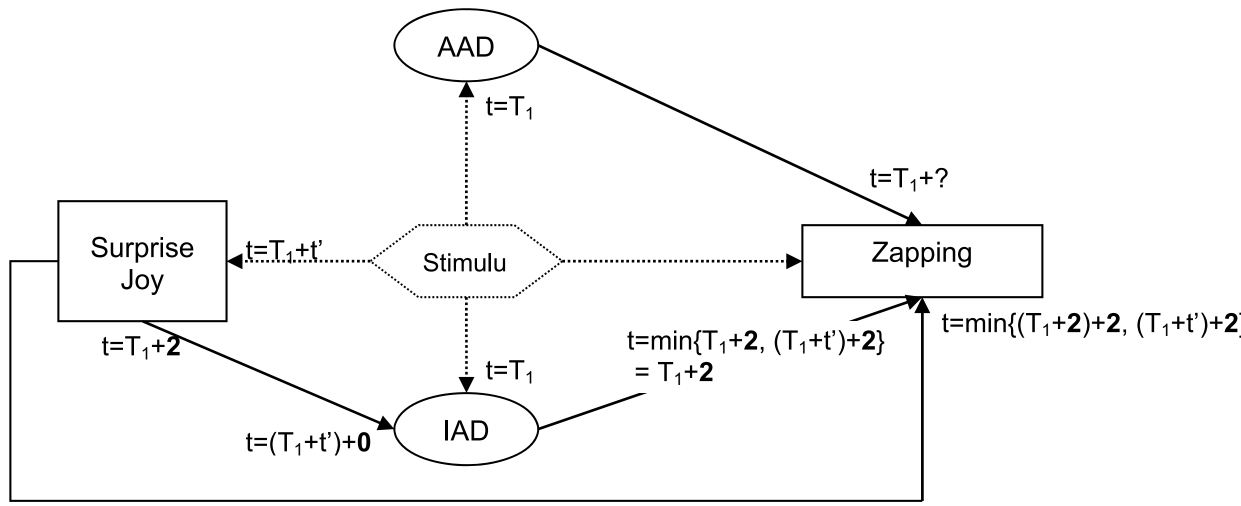
FIGURE B.1: DEFINING LATENCY RESPONSE BASED ON OPTIMAL LAG EFFECTS



Similarly, it becomes important to understand the latency of the zapping response from changes in emotions and IAD. So, we proceed with binary (Probit) regressions of zapping on emotions, IAD, AAD and all other individual, commercial and brand-specific covariates, altering the lag variables independently for surprise/joy and IAD. Evidently, these emotional-based effects can not cause physical hand movement reactions such as zapping simultaneously so we only test for values of lag corresponding to 250 to 1500 msec. Figure B.1 (right graph) shows that the predicted positive effects of both surprise and IAD on zapping occur more strongly with a lag of 2 frames for each. Thus, the final model should incorporate direct effects of these constructs with latency from 500 to 750 msec., as the appropriate mean peak response.

Figure B.2 summarizes the results. It depicts a flow model of latency effects (i.e. time measured only in flows) from stimulus onset to the time at which a construct causes an effect. It shows the latency responses in bold that are used in the final model.

FIGURE B.2: RESPONSE LATENCIES FOR EMOTIONS, ATTENTION AND ZAPPING



Note- Numbers in bold are the result of lag regression estimates in Figure B.1; t' is the latency from visual to facial response; t'' is latency from visual to physical response. For a certain individual, IAD influences AAD only when the number N of viewers is small. In this case, N=50 makes the relation practically negligible.

References

Hansen, Christine H. and Ranald D. Hansen (1994), "Automatic Emotion: Attention and Facial Efference," in *The Heart's Eye: Emotional Influences in Perception and Attention*, P.M. Niedenthal and S. Kitayama, eds. San Diego: Academic Press, 217-43.

Nummenmaa, Lauri, Jukka Hyönä and Manuel G. Calvo (2009), "Emotional Scene Content Drives the Saccade Generation System Reflexively," *Journal of Experimental Psychology: Human Perception and Performance,* 35(2), 305-23.