ABSTRACT

Title of dissertation: STATISTICAL ANALYSIS OF ONLINE EYE AND FACE-TRACKING APPLICATIONS IN MARKETING

Xuan Liu, Doctor of Philosophy, 2015

Dissertation directed by: Professor Michel Wedel
Department of Marketing
Robert H. Smith School of Business

Eye-tracking and face-tracking technology have been widely adopted to study viewers' attention and emotional response. In the dissertation, we apply these two technologies to investigate effective online contents that are designed to attract and direct attention and engage viewers emotional responses.

In the first part of the dissertation, we conduct a series of experiments that use eye-tracking technology to explore how online models' facial cues affect users' attention on static e-commerce websites. The joint effects of two facial cues, gaze direction and facial expression on attention, are estimated by Bayesian ANOVA, allowing various distributional assumptions. We also consider the similarities and differences in the effects of facial cues among American and Chinese consumers. This study offers insights on how to attract and retain customers' attentions for advertisers that use static advertisement on various websites or ad networks.

In the second part of the dissertation, we conduct a face-tracking study where we investigate the relation between experiment participants' emotional responses

while watching comedy movie trailers and their watching intentions to the actual movies. Viewers' facial expressions are collected in real-time and converted to emotional responses with algorithms based on facial coding system. To analyze the data, we propose to use a joint modeling method that link viewers' longitudinal emotion measurements and their watching intentions. This research provides recommendations to filmmakers on how to improve the effectiveness of movie trailers, and how to boost audiences' desire to watch the movies.

Statistical Analysis of Online Eye and Face-tracking Applications in Marketing

by

Xuan Liu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor Michel Wedel, Chair/Advisor
Professor Paul Smith
Professor P.K. Kannan
Professor Yuan Liao
Professor Kasso A. Okoudjou

Dedication

To my loving and supportive husband, Huashuai Qu.

Acknowledgments

I would like to express my appreciation to everyone who made it possible for me to complete this dissertation.

First and foremost, I own my deepest gratitude to my advisor, Professor Michel Wedel for his continuous support of my Ph.D research, his excellent guidance, caring, patience, enthusiasm, and immense knowledge. Without his persistent help, this dissertation would not have been possible. I am very fortunate to have him as my advisor.

Besides my advisor, I would like to thanks Professor Thales S. Teixeira, from Harvard Business school for generously sharing his time discussing and brain storming on the comedy movie trailer project. I am thankful for his aspiring guidance and invaluable advice.

I am also grateful to Professor Qiuzhen Wang, from Zhejiang University for spending countless hours in data collection, discussing ideas and writing in the eye-tracking study of commercials.

My sincere thanks also goes to the rest of my thesis committee members: Professor Paul Smith, Professor P.K. Kannan, Professor Yuan Liao and Professor Kasso A. Okoudjou, for their encouragement, guidance and insightful feedbacks.

I would also like to thank everyone in the department of Mathematics. I could not complete my work with my amazing friends here. I'm grateful to Yu Guo, Marie Chau, Xia Li, Xuan Yao, Chen Dong and Wen Zhou for their friendship and support in every moment of my doctoral journey.

I own my deepest thanks to my parents and my dear friends: Renting Xu, Maggie Guo, Ying Wang and Ying He, who stood by me through all good times and bad. Special thanks to my loving, supportive and encouraging husband Huashuai, who is always there for me. I own them everything and wish I could show them how much I love and appreciate them.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Effective online content is designed to attract viewers' attention and engage their emotional response. Modern technologies such as eye-tracking and face-tracking enable advertisers and marketing researchers to measure consumers' attention and emotion during exposure to commercial stimuli in a nonintrusive way and help them gain insights for producing effective advertisements.

1.1 Eye-tracking and Face Tracking

Eye-tracking

Eye-tracking technology has been widely adopted in various disciplines including psychology, cognitive science, medical research and marketing research. When visual information is being processed, the human brain directs the eyes to the area that contains the information. Therefore, we can observe the path of the observer's attention by tracking his or her eye movements.

Eye movements include fixations and saccades. Fixations happen involuntarily when eye gaze is focused on a certain location. During fixation, neurons in the early visual areas of the brain are continuously stimulated to maintain visibility of the stimuli and the fovea delivers the visual information. Fixations on average last 200 milliseconds for linguistic text and 350 milliseconds for a scene. Saccades are the rapid eye movements from location to location, which can be both voluntary and involuntary. When a saccade happens, the eyes move as fast as they can. The speed of the movement cannot be controlled consciously. It usually last 200 milliseconds to prepare a saccade to the next location. A series of fixations and saccades form a scanpath. Information about the stimuli is obtained mostly during the fixations rather than saccades. Therefore, to understand what information has been processed, we can examine the locations of fixations in a scanpath.

An eye tracker is a device that measures eye positions, gaze directions and eye movements. There are several types of eye trackers: Some measure eye movements via a special contact lens attached to the observer's eyes that records the movement of the lens; Some measure electric potentials by electrodes placed around the eyes. Modern eye trackers use a much less obtrusive and cost effective method called *pupil center corneal reflection eye tracking* by applying fundamental principle of corneal reflection (CR). Figure 1.1[1] is an illustration of an eye tracker from Tobii, one of the world leading eye tracking companies. The eye tracker is installed at the bottom of the computer scene. During the eye-tracking experiment, the eye tracker emits infrared/near-infrared light to create corneal reflection. A camera on the computer captures the image of viewers' eyes. The eye-tracking software identifies the darkest region in the eye as the pupil, and the lightest spot as the corneal reflection. As gaze direction changes, the relative positions between the center of the pupil and

[1] http://www.tobii.com/en/about/what-is-eye-tracking/

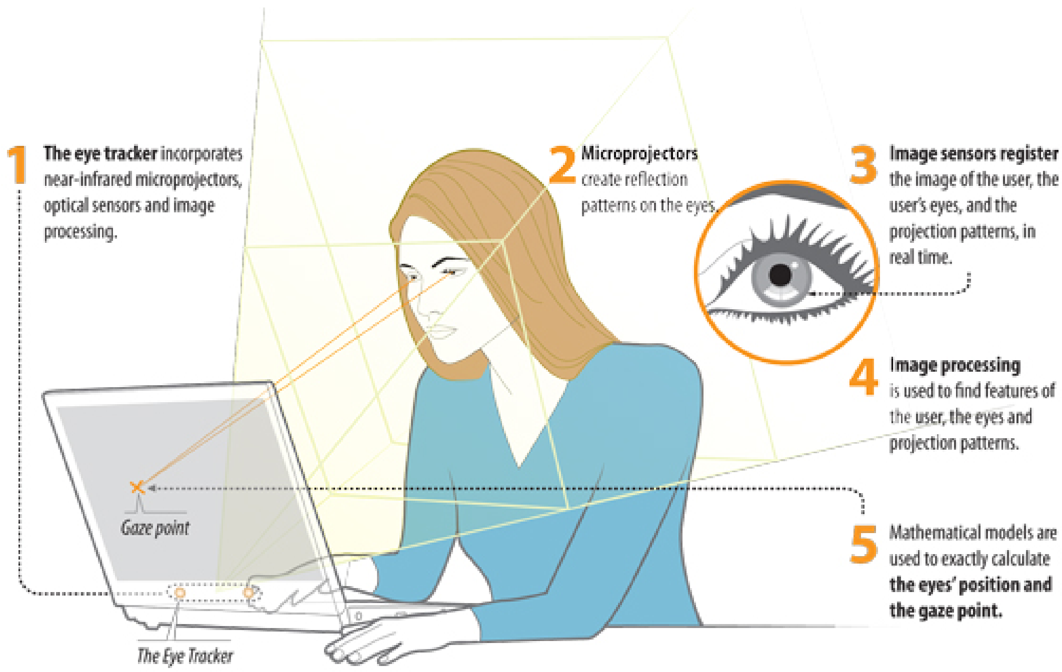


Figure 1.1: Eye-tracking system (Tobii)

the corneal reflection change accordingly. The location of these two are captured to compute the fixation point on the computer scene. Other measurements such as pupil size and pupil dilation, which could be indicators of excitement, can also be measured during the eye tracking process.

In marketing research, usage of eye-tracking technology has grown rapidly in the last decade. Advertisers and researchers use it to getting insights for optimizing advertisements. For example, Wedel and Pieters [2008, 2014][1, 2] investigate the effectiveness of advertisements by examining how viewers distribute their attention as they scan the online commercial contents.

Face-tracking

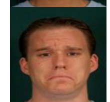
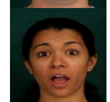
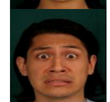
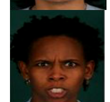
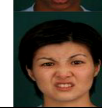
Facial expressions are strongly associated with emotional state. Face-tracking technology analyzes viewers' facial expressions and identifies their emotions using the Facial Action Coding System (FACS) developed by Ekman and Friesen (1987). Through facial movements, FACS can anatomically measure facial expressions of a human being.

The fundamental component of observable facial movements is called an Action Unit (AU), which contains a contraction or relaxation of one or more muscles. A total of 46 AUs were defined by Ekman and Friesen to describe the independent facial muscle movements. They also defined a few Action Descriptors (ADs), which involve actions of several muscle groups. Letter A-E were used as the intensity scores of the AUs (A: Trace of the action, B: Slight evidence, C: Marked or pronounced, D: Severe or extreme and E: Maximum evidence). For example, AU 5B means there is slight evidence of Action Unit 5. Table 1.1 shows AUs for the 6 basic emotions (Happiness, Sadness, Surprise, Fear, Anger and Disgust) ([3]). The description of AUs involved in these emotions and their underlying facial muscles are described in Table 1.2[2].

FACS is a common standard to recognize emotions. Ekman and Friesen published the FACS manual in 1978, which describes in detail how to categorize facial expressions based on facial muscle actions. The manual guides human coders to learn the technique to deconstruct facial muscle movement into AUs and manually

[2] http://www.cs.cmu.edu/ face/facs.htm

Table 1.1: Facial actions for six basic emotions (FACS Manual [3])

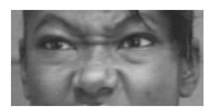
Emotion	Action Units	Example Image
Happiness	6 + 12	
Sadness	1 + 4 + 15	
Surprise	1 + 2 + 5B + 26	
Fear	1 + 2 + 4 + 5 + 7+ 20 + 26	
Anger	4+ 5+ 7+ 23	
Disgust	9 + 15 + 16	

identify emotions based on the combinations of AUs. To identify an emotion, at least two independent certified FACS encoders have to agree on a conclusion, for subjectivity and accuracy issues. The training and manual coding process could be very time consuming and not cost efficient.

Novel technologies provide researchers facial imaging systems that automatically detect and produce emotion profiles from faces in videos ([4, 5]). Specialized emotion recognition software makes it possible to process real-time video data at a rate of four images per second([1]). The advantages of facial imaging using artificial intelligence include consistency, scalability and repeatability of the measurements.

Figure 1.2 provided by nViso, a facial imaging technology company, shows the mechanism of how the facial imaging system works. The emotion recognition software superimposes a "spider web" mask that contains 143 points to linked to the

Table 1.2: List of Action Units and the underlying facial muscles (FACS Manual)

AU	Description	Facial Muscle	Example Image
1	Inner Brow Raiser	Frontalis, pars medialis	
2	Outer Brow Raiser	Frontalis, pars lateralis	
4	Brow Lowerer	Corrugator supercilii, Depressor supercilii	
5	Upper Lid Raiser	Levator palpebrae superioris	
6	Cheek Raiser	Orbicularis oculi, pars orbitalis	
7	Lid Tightener	Orbicularis oculi, pars palpebralis	
9	Nose Wrinkler	Levator labii superioris alaquae nasi	
12	Lip Corner Puller	Zygomaticus major	
15	Lip Corner Depressor	Depressor anguli oris (a.k.a. Triangularis)	
20	Lip stretcher	Risorius w/ platysma	
23	Lip Tightener	Orbicularis oris	
26	Jaw Drop	Masseter, relaxed Temporalis and internal Pterygoid	

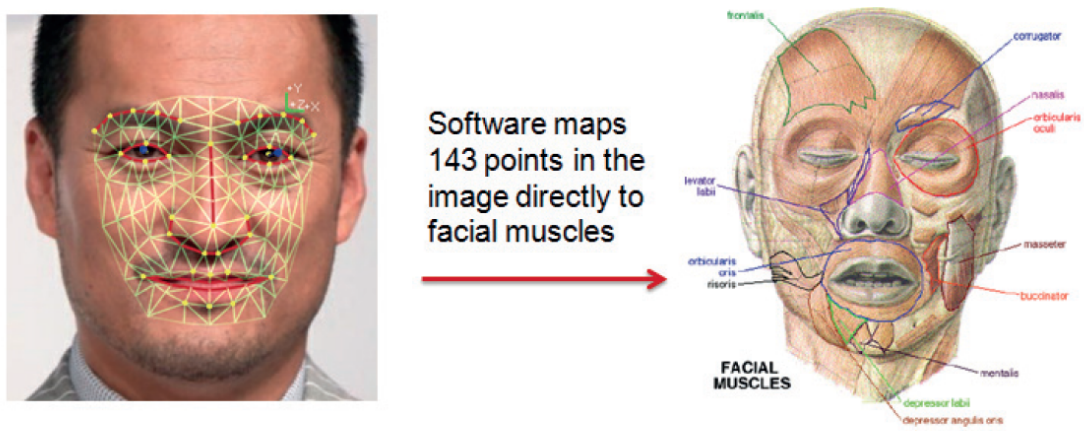


Figure 1.2: Facial imaging mechanism (nViso)

key muscles on a human face. Facial muscle movements result in location changes in these points on the mask which are captured in real-time by the computer. Machine-learning tools are then applied to analyze the facial actions in response to stimuli and categorize them into the six basic emotions based on FACS ([6]).

Due to the essential role emotions play in our daily life, face tracking has a broad range of applications in many areas including education, behavioral science, mental health and deception detection. It has also been applied to marketing research as a robust and nonintrusive tool to measure consumer responses to various marketing stimuli, such as static advertisement and video commercials.

This thesis includes two studies. The first is an eye-tracking study of static e-commerce websites to investigate the joint effect of facial expression and gaze direction. The other is a face-tracking study of the impact of emotional responses to online video content.

1.2 Bayesian Statistical Models for Eye and Face Tracking Data

In this dissertation, we apply Bayesian statistical models to analyze eye and face tracking data. Bayesian statistics assumes parameters to be random variables that have their own distributions. It utilizes Bayes' rule to estimate posterior parameter distributions given observed data. Instead of numerical integration to obtain posterior distributions, the Markov Chain Monte Carlo (MCMC) method provides an alternative way of drawing samples from posterior distributions. Software, such as BUGS and JAGS have been developed to make the implementation of MCMC much easier.

To increase accuracy, eye tracking experiments often use repeated measurement designs. Data hierarchies need to be considered. Missing data, due to failure of calibration, interruption of the experiment and measurement error, can also cause potential problems in classical ANOVA methods. What's more, data collected from eye-tracking studies usually have varying distribution properties. For example, three main responses are measured in a typical eye-tracking study: (1) Fixation count (measuring attention retention) in each AOI: The count data follows a Poisson distribution. However, if an AOI have a higher probability of not being fixated on, the distribution of fixation count will be zero-inflated. (2) Fixation duration (measuring depth of processing): Duration time is a continuous non-negative measurement. (3) Time to first fixation (measuring attention selection): when an AOI is never fixated on, time to first fixation will not be recorded or will be right censored to the total dwell time on the image. Adopting Hierarchical Bayesian ANOVA model solves all these problems in eye-tracking experiment at once. Hyper-priors are imposed to account for individual heterogeneity. Missing data are treated as extra parameters. We can also specify the special distributions for different dependent variables.

Face-tracking experiments is marketing research usually generates two types of data: emotion data collected while participants watching video contents, and binary or multinomial data related to participants purchase intentions or rating data collected from questionnaires. There are many classical approaches for analyzing these data separately, including linear mixed effects models for the emotion data, and generalized linear mixed effects models for the intention data. However, it might be more appropriate to use a joint model for making statistical inferences, as participants emotion while watching video contents is likely to correlated with purchase intentions. The key in the joint model is to connect these two separated models through a latent Gaussian process. Traditional maximum likelihood estimates can be obtained via the EM algorithm, but the implementation is rather complicated. The Bayesian approach to the joint model enables us to make full posterior inference for the parameters, and the implementation of the approach in BUGS or JAGS in fairly clear. In addition, when it comes to comparing different structures for the latent Gaussian process, the Bayesian approach with the DIC criterion makes it easy to test different structures without the need to change the likelihood function every time.

1.3 Facial Expression and Gaze Direction in Commercials

The booming of global e-commerce has brought a growing challenge to online retailers: in the face of increasing competition and decreasing consumer attention spans they need to attract customers to their websites and entice them to look at their products and brands. To accomplish this, e-tailers have focused on website design elements that attract and retain consumers attention. Human touch factors, in particular human models, have been popular and appear in a wide range of e-commerce websites, prints, video and TV ads, including those for apparel, perfumes, cosmetics and accessories. Images of models have been demonstrated to attract visual attention ([7]), improve the perceived social presence ([8]), and promote trust ([9]). Especially the human face has been shown to capture attention more readily than other stimuli ([10, 11]), and is often used as a means to attract attention to products sold on websites. Indeed, the presence of faces in print advertisements has been found to positively affect attitudes towards the ad and the brand, as well as purchase intentions ([12]). Moreover, facial cues may enhance the impact of human faces on the viewer. Gaze direction and the facial expression of emotion are the most important facial cues, and are the two key dimensions of non-verbal communication ([13]). They are also the focus of the present research.

First, a persons gaze direction communicates her current focus of attention and provides clues about which objects are important to her, and about her interests and intentions towards them ([14, 15]). This has been shown to cause viewers to automatically orient their attention to the target of another persons gaze ([16]). As a consequence, gaze cueing is finding its way into marketing practice as a tool to direct attention. Website designers have capitalized on the tendency of viewers to orient towards the product that a model on a website is looking at. However, research on the effectiveness of this practice is scarce. In the only study in marketing to date, Hutton and Nolte (2011) [17] demonstrated that gaze cueing through simultaneous eye, head and body position had a positive effect on attention to print ads and the products shown in them.

Second, individuals facial expressions are a reliable indication of their emotions ([18]). On websites, the use of a models facial expression capitalizes on the idea that the emotion depicted may carry over to the product shown on the site. Despite the extensive research on facial expressions in psychology, there have been only a few studies looking at their effect in marketing. For example, Cho and Norbert (2006) [19] found that the preference for (ones own) smiling face on a website can cause a higher liking of the product displayed on that site, and Small and Verrochi [20] showed that the use of emotional expressions in charity advertising positively affects consumers attitudes towards the ads in question.

In our first study, we are interested in the joint effects of these two basic facial cues on attention. Intuitively, one could expect that, for example, a happy face looking at a product would have a stronger effect on directing viewers attention than a face with a neutral expression. It would provide information about not only the location of the product of interest, but also about its emotional significance: a smiling face looking at a product would inform the observer that the model is looking at something enjoyable. However, although basic research has addressed this question, the verdict on whether the two facial cues have this joint effect is still out ([21, 22]). In addition, there has been little research effort directed at systematically exploring the joint impact of these two facial cues in naturalistic contexts. Yet this is important, because while in most prior research faces are shown in isolation, in naturalistic contexts such as websites they are shown against a background scene along with other objects that compete for viewers attention, which may reinforce or inhibit their impact. This study aims to fill that gap in our knowledge by examining the joint attention effect of gaze direction and facial expression of models on e-commerce websites.

While such research has obvious implications for the design of e-commerce websites, with the global reach of most of these websites it becomes critical to understand commonalities and differences in the perception of facial cues between cultures. Many e-tailers adapt their websites to the local culture, not only through the language and imagery of product descriptions, but also through the models displayed and their facial cues. Indeed, it has been shown that there are differences in the eye movements made by people in different cultures while viewing a human face ([23]), and that people's perception of emotional expressions is influenced by their cultural background ([24]). Also, perceived ethnic similarity of the model and the viewer has been shown to result in more positive evaluations ([25]). In the first study we use eye tracking to investigate the joint effects of facial expression and gaze direction on eye movements among American and Chinese consumers. In addition, we study the effect of the ethnic match of the viewer and the model.

This study intends to make the following contributions. First, it empirically explores the joint effects of facial expression and gaze direction of models on a wide range of indicators of consumers attention in the naturalistic context provided by e-tailer websites. Second, it investigates similarities and differences in the effects of facial cues among American and Chinese consumers. Third, it investigates the effects of the ethnicity of the model, and the extent to which it moderates the effect of the facial expression of the model. This research thus intends to offer new insights into the effects of facial cues with the aim of providing recommendations for website design on how to take advantage of them to attract and retain customers attention. In the next section we provide a review of the relevant literature. The details of three eye tracking experiments are described in the following sections. The final section provides a general discussion of the results and the implications, and future research directions.

1.4 The Impact of Emotional Responses in Movie Trailers

Movies are of great importance in global economy. According to the 2012 theatrical statistics summary conducted by the Motion Picture Association of America Inc. (MPAA), global box offices across the world are $34.7 billion in 2012 alone. The United States has one of the oldest film industries and is one of the largest markets for movies by box office. The domestic box office in 2012 was $10.8 billion. In 2013, 225 million people, 68% of the U.S./Canada population (Ages 2+), went at least once to a movie theater. About 13% percent of these go to the theatre at least once per month.

In the motion picture industry, the success of a movie is highly unpredictable. To reduce the uncertainty by promoting movies, the most popular marketing tool used to advertise a film is the movie trailer, or movie preview. A movie trailer usually has time length up to 2 minutes and 30 seconds, as regulated by the MPAA. However, there is no restriction on the duration of Internet or home video trailers. Movie trailers have been shown to be the most influential factor on the intentions to watch a movie ([26]).

National advertising campaigns of movies are aired in theatres, on television, on the Internet and in various home video formats, weeks or months prior to the release of the movie. Distributors usually pay the cost of movie advertising and the budget is normally set as a fixed percentage of the movies production cost ([27]). The cost of making a movie trailer therefore ranges from $300,000 to $600,000 ([28]). According to PricewaterhouseCoopers (PwC)s 2013 entertainment report[3], cinema advertising spending was $741 million in U.S and $2.6 billion worldwide in 2012 alone ([29]).

Effective video movie trailers are designed to induce desirable emotional responses with the purpose of making consumers want to watch the movie. Our research uses web-based face-tracking to study viewers real-time emotional responses when watching online video trailers and links them with viewers intentions after watching the contents. Since watching intention includes not only theater tickets purchase intention but also home DVD or blue-ray purchase and renting intention, it is a reasonable and comprehensive measure in our study of movie trailers, and one

[3]PwC is a multinational professional serves network

that is widely used in practice. The aim is to predict these measures of intention to watch the movie in order to facilitate optimization of trailers. For this purpose, a joint hierarchical Bayesian model is developed, which consists of linked sub-models for the moment-to-moment emotional responses and the intention of watching a movie, respectively.

To study the continuous moment-to-moment (MTM) emotion response data, previous research has mainly focused on the effects of the aggregated emotions, such as average emotion intensity, linear trends of changes in emotions, and peak and end emotion levels ([30]). Elpers et al. [2002] [31] examined the effects of weighted averages of emotions of a sample of participants obtained by functional data analysis of zapping behavior of TV commercials. Their results show that higher levels and higher velocities of pleasantness decrease the zapping rate for TV commercials. Hui et al. [2014] [32] also adopted functional linear models and revealed that the last quantile of the contents of a TV show is weighted more than the first quantile, while peak and trough patterns do not have significant effects on overall evaluation of the show.

In our study, we develop a new joint model to simultaneously examine the emotion changes over time and the relationship between emotion and the overall decision. Joint models have been widely used in the areas of medical statistics to analyze longitudinal and time to event data. In general, there are two sub-models combined in a joint model, one for the longitudinal data and the other for end-point data. These two sub-models are connected through one or more subject-specific random effects ([33]). Prior applications involve those by Tsiatis and Davidian [2004]

[34], and Wang et al. [2012] [35]. A joint model allows inferences for three aspects simultaneously: the effect of covariates on longitudinal processes, the association between the longitudinal process and end-point variables and the effects of the covariates on the end-point variables. The covariate effects on end-point measures thus include two parts: the direct covariate effect on the end-point measure and the indirect treatment effect on end-point measures through the latent longitudinal process.

1.5 Outline of the Dissertation

In this dissertation, we investigate the problem of how to design effective online contents, including static e-commerce website and online video contents. Eye-tracking and face-tracking tracking technologies are applied in these two studies to examine viewers' attention and emotion engagement. The findings in the first study can help e-tailers to design websites among viewers from different cultural background to efficiently direct their attention to the product or brand. The second study provides insights to filmmakers in how to place the emotion-inducing contents in comedy movie trailers to attract the audiences to watch the movie.

Chapter 2 describes three eye tracking experiment to investigate the joint effects of facial expression (neutral/happy) and gaze direction (direct/averted) of models on websites on visual attention among American and Chinese participants. They reveal that among both cultures a gaze cue primes initial attention to the product or brand and show that positive affect from the happy expression when a model looks at the viewer carries over to the product or brand. For American participants, a model that looks at the viewer with a happy expression draws more attention to the brand, while for Chinese participants a model that looks at the product with a happy expression draws more attention to the brand. These differences are explained from a cultural difference in using the eyes and mouth as cues to recognize and interpret smiles in Asian and Western cultures, respectively. Further, the racial match between a model and the viewer exacerbated the attention effects of facial expression.

Chapter 3 proposes a joint statistical model to study the longitudinal emotion data collected online while viewers are watching a movie trailer and the viewers' intention to watch the movie. We calibrate the model using experimental data collected from participants facial responses while watching comedy movie trailers on the web and relate diagnostic emotion-inducing scenes to the viewers decision of watching the movie afterwards. Our findings show a significant positive effect of joy-inducing content at the peak and end scene on watching intention. Negative emotion-inducing content placed at the beginning of the trailer has a negative effect. High music volume at the peak and low volume at the beginning also increase watching intention. The proposed method and results allow filmmakers to optimize the effectiveness of movie trailers, so that they can increase peoples desire to watch the movie by balancing the amount of highly emotion-inducing content shown in order to maximize interest, while not giving away too much of the story in the trailer.

Chapter 4 concludes our findings in the e-commerce websites and comedy movie trailer studies. It summarizes the contributions and makes suggestions to online retailers and filmmakers based on results from these two studies. It also points out the limitations of these studies and directions of future research.

Chapter 2: Facial Cues, Gaze Direction and Attention Patterns in Different Cultures

2.1 Introduction

Gaze Direction

In recent years, the perception of eye gaze direction has emerged as critical in face processing ([36]). For example, Senju and Hasegawa (2005) [37] showed that the detection of a target in the periphery of vision is impeded when observers fixate on a face that is looking directly at them, compared to fixating a face with averted eyes. Thus, it appears that direct eye gaze captures and retains attention to the face (see also Bindemann et al. 2005 [38]). Averted eyes, on the other hand, have the ability to shift a viewers attention and lead to a faster classification of targets in the direction of the perceived gaze (e.g., Driver et al. 1999; Friesen, Moore and Kingstone 2005; Friesen and Kingstone 1998; Kingstone et al. 2004; Langton, Watt and Bruce I 2000 [39, 40, 16, 41, 42]). Research has revealed specialized regions in the human visual brain that are responsive to eyes, eye movements, and gaze direction (Superior Temporal Sulcus; STS), some neurons in these regions being sensitive to orientations of body, head and eyes simultaneously, others to only one of these ([14]). Hutton and Nolte (2011) [17] investigated simultaneous eye, head and body direction cues of a model in a print advertisement and found that participants looked longer at the product and brand regions of the advertisement as well as at the entire advertisement when the models gaze was directed towards the product rather than towards the viewer.

Based on these findings, we hypothesize that the gaze direction of models on websites promotes customers attention to a product present in the direction of the perceived gaze, and that this increased attention may spill over to the brand and the text containing the product description.

Facial Expression

A facial expression is caused by contraction of facial muscles and conveys the emotions of an individual to the viewer. The amygdala plays a central role in processing these facial expressions of emotions ([22]). The so called "contagion theory" postulates that viewers automatically mimic facial expressions, and that proprioceptive feedback from this behavior affects the emotional experience of the viewer ([43, 44]). Research has confirmed that individuals spontaneously mimic emotional facial expressions in static pictures and experience subsequent emotional contagion ([45, 46, 43, 47]), even when they are unconsciously exposed to these facial expressions ([46]). Moreover, it was shown that the contagion effect causes facial expressions to attract attention to the face ([48]). Angry faces have the capability to inhibit attention disengagement causing longer gaze, but the effect for happy faces is less strong ([49, 22]). In the present research we focus on expression of happy emotions, because these are by far the most frequently used on e-commerce websites. Based on the research reviewed above, we hypothesize that the facial expression of models on websites positively impacts consumers arousal, and that this will positively affect attention to the face of the model, which carries over to the product/brand.

The Joint Effect of Facial Expression and Gaze Direction

Gaze direction plays a critical role not only in orienting viewers attention but also in viewers perception of emotions. On the one hand, Coss, Marks and Ramakrishnan (2002) [50] argued that staring can be perceived as aggressive, and Nichols and Champness (1971) [51] accordingly found that it led to increased galvanic skin response compared to averted eye gaze. On the other hand, it has been argued that gaze direction can act as a signal of attraction between people. But there is some disagreement about this. While Mason, Tatkow and Macrae (2005) [15] found no difference, Ewing, Rhodes and Pellicano (2010) [52] found that people find direct gaze more attractive. Schilbach et al. (2006) [53] found that participants felt more engaged with a virtual agent that looked at them directly, as compared to one that looked at another person. Thus, there is no unique view on the effect of gaze direction on viewers emotional experience. In addition, the emotional experience caused by gaze direction cues may interact with that caused by facial expression cues.

Research has shown that gaze direction and facial expression have interactive effects on emotion recognition ([21, 54]), perceived attractiveness of faces ([55, 56]), object preference ([57]), and attention orienting ([58]). Specific areas in the visual cortex (STC and amygdala) have been shown to play a role in the analysis of both facial expression and gaze direction ([22]). This makes interactive effects of facial cues possible, although it has also been argued that facial cues are processed independently ([22]). For example, researchers found that angry faces are recognized as expressing more anger with a direct than with an averted gaze, whereas fearful faces are recognized as expressing more fear with an averted than with a direct gaze ([54]). Adams and Kleck (2003) [21] provided an overarching framework explaining these effects: approach-oriented emotions (anger and happiness) are better recognized with direct gaze, and avoidance-oriented emotions (fear and sadness) are better recognized with averted gaze. Jones et al. (2006) [56] found that faces with direct gaze are perceived as more attractive when smiling than when holding a neutral expression, whereas faces with averted gaze are less attractive when smiling than when holding a neutral expression (see also Conway et al. 2008 [55]).

Importantly, Bayliss et al. (2007) [57] reported that the objects gazed at with a happy expression were liked more than objects one gazed at with a disgusted expression, but objects that were not gazed at were liked equally for both expressions. In addition, the results of Ozono, Watabe and Yoshikawa (2012) [58] support the hypothesis about avoidance behavior: when someone is observed looking in a particular direction with a fearful expression, the observer tends to avoid that direction. But they did not find evidence to support approach behavior: when someone is observed looking in a particular direction with a happy expression, the observer would not tend to look in that direction more.

Thus, the findings on the joint effect of gaze direction and facial expression are still somewhat inconclusive. In line with this, six experiments, Hietanen and Leppanen (2003) [22] did not show evidence to show that facial expression affects the attention orienting triggered by gaze direction. Contrary to the conclusions of Adams and Kleck (2003) [21], Bindemann, Burton and Langton (2008) [36] found that averted eye gaze slows the categorization of any facial expression, rather than having a selective impact on specific emotions. They suggested that gaze direction affects the analysis of facial expression via an intermediate process, namely the allocation of visual attention to the target face. Facial expressions might only affect attention (dis)engagement, as argued by Fox, Russo and Dutton (2002) [49]. Thus, the verdict on whether these two facial cues have a joint effect is still out, and there is a need to further investigate the joint attention effects of gaze direction and facial expression.

Culture Differences in Effects of Gaze Direction and Facial Expression

Facial expressions have long been considered the universal language of emotion, being the same across cultures ([59, 60]). Yet, there is evidence that the perception and recognition of emotional expressions is influenced by the cultural background of the viewer. Importantly, it has been shown that Western participants weigh facial expression cues displayed in the mouth more when judging emotions, whereas

Asian participants have been found tend to weigh cues in the eyes more ([61]). Indeed, using eye-tracking technology, Jack et al. (2009) [62] reported that Western participants distributed their fixations evenly across the face, whereas Asian participants persistently fixated the eye region more (see also Blais et al. 2008 [23]. Further, Jack, Caldara and Schyns (2012) [24] revealed cultural differences in the expectations of facial expressions: Westerners expectations predominantly involve the mouth, whereas Asians show a preference for expressive information in the eye region.

The smile is the most common and universal human facial expression ([59]), and this is the facial expression we focus on in this study. Smiles can be classified as Duchenne or non-Duchenne. Non-Duchenne smiles involve only the (zygomatic major) muscles around the mouth area, while Duchenne smiles also involve the (orbicularis oculi) muscles around the eyes ([63]). The interpretation of the meaning of any of these two smiles depends on whether eyes or mouth are used as a source of information ([64]). In Western cultures, the mouth is most crucial diagnostic feature in identifying and interpreting smiles ([65, 66, 67, 68]), but Asians rely to a larger extent on information from the eyes ([64]). In line with this, Ozono et al. (2010) [69] found that Asian participants found faces with greater upper-half (eyes) smile intensity more trustworthy, but American participants found faces with greater lower-half (mouth) smile intensity more trustworthy.

Although no research has yet looked at cultural differences in the effects of gaze cues on attention, several studies have found differences between cultures in gaze behaviors during face-to-face communication. Bond and Komai (1976) [70] reported that Asians felt uneasy when being gazed at directly. Consistent with this, Pierson and Bond (1982) [71] found that Chinese conversational partners dont usually look at directly at each others eyes, whereas Westerners do ([72]). These studies support the idea that perceived gaze direction may have different effects in Asian and Western cultures.

Based on the above research we hypothesize that the effect of facial cues on attention patterns will be different in different cultures. Asians, paying more attention to the eyes, may be more sensitive to gaze direction and show a more prominent gaze following effect, while Westerners, paying more attention to the mouth, may be more sensitive to emotional expressions of happiness. The latter also suggests that facial expressions moderate the effect of gaze cues differently in Western and Asian cultures.

Ethnic Similarity Effects

Ethnicity is an integral part of consumers' perceptions of a spokesperson or model (Huston, D'Ouville and Willis 2003) ([73]). Theories of similarity attraction and in-group bias have been used to explain the role of ethnicity in social interactions. Similarity attraction theory posits that a users perception of similarity with another on attributes such as ethnicity increases interpersonal attraction and liking ([74]). In-group bias theory ([75]) proposes that people behave more favorably towards in- group than towards out-group members. Ethnicity, as a readily observable physical cue, is readily used to classify other people as either in-group or out-group

([76]).

There has been extensive research examining the effect of the ethnicity of models displayed in advertising. This research has demonstrated that perceived ethnic similarity results in more positive ad evaluation ([77, 25, 78]), higher perceived trustworthiness of spokespersons ([79]), and higher likelihood of purchase ([80, 81]). Recently, a few studies have shown that a match of the ethnicity of viewers and online models positively affects the viewers evaluative responses in e-commerce settings ([82, 83]). Kareklas and Polonsky (2010) [84] revealed that model-viewer ethnic similarity is most important for Hispanic participants, followed by black participants, then by Asian participants, and least so for white participants.

In addition, it has been shown that the perception of emotional expressions can be influenced by the perceivers feeling of belonging to the same (in-group) or a different ethnic group (out-group) ([85]). Brown, Bradley and Lang (2006) [86] found that people had stronger affective responses when viewing pictures of ethnic in-group as compared to ethnic out-group members. Krmer et al. (2013) [85] found that when a model gazed directly at the viewer, participants from the ethnic out-group tended to perceive emotions as more pronounced than participants from the ethnic in-group. Based on this literature, we hypothesize that the ethnic match between a model and the viewer not only positively affects attention to the face and product information, but also moderates the effect of the models facial expression to become stronger.

This Research

We test our predictions in three eye-tracking experiments. In the first experiment, we manipulate the gaze direction (direct/averted) and the facial expression (neutral/happy) of models on e-commerce websites in a study among American participants, and investigate the effects on multiple indicators of attention to the face, eye, product, brand logo, and text, as well as the webpage as a whole. In the second experiment, we replicate this study among Chinese participants. We focus on Chinese and Americans, because these are some of the worlds largest populations. Because we need to adapt the details of the experimental designs to the local conditions in the United States and China, we analyze and report these two experiments separately, rather than jointly. In the third experiment, we manipulate the facial expression (neutral/happy) and the ethnicity (Caucasian/Asian) of models on e-commerce websites in a study among Chinese participants, to examine the effects of viewer-model ethnicity match on several indicators of attention to the webpage as a whole and its various elements. In these three experiments, while gaze direction is manipulated through the eyes of the model, facial expression is manipulated through the mouth. This enables us to investigate differences in the effects of these two facial cues between American and Chinese participants.

2.2 Experiment with American Participants and Caucasian Models

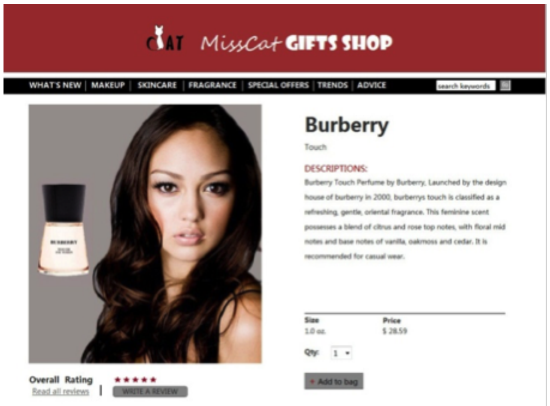
This experiment used a 2×2 between-subject design, with two levels of facial expression (neutral versus happy) and two levels of gaze direction (direct: looking straight ahead versus averted: looking at the product). The experimental conditions were obtained through digital manipulation of images of faces. This allows us to control for other features of the face, by changing only very specific facial cues in the eyes and mouth.

Stimuli

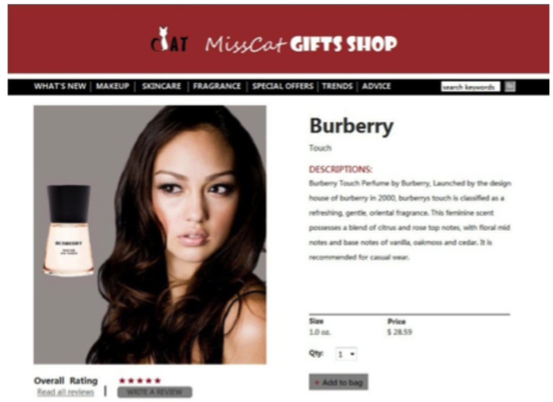
We obtained from the internet 22 images of female Caucasian fashion models shown above the shoulder, facing the camera, and with a neutral facial expression. Adobe Photoshop digital imaging software was used to manipulate the mouth and eyes in the original images, to produce happy faces or faces with averted gaze. In a pretest, 20 participants were asked to rate the images on attractiveness, naturalness, and feelings about the smile. Based on the results 16 images were chosen as stimuli. Manipulation checks among 126 participants in the main study (see below) revealed a significant difference ($t = 2.013$, $p = 0.046$) in perceived smile between the images with a happy face (mean = 4.11, SD = 1.78) and a neutral face ($mean = 3.44$, $SD = 1.93$), and a significant difference ($t = 6.77$, $p = 0.000$) in perceived gaze direction between images with a direct gaze (mean = 5.66, SD = 1.39) and those with averted gaze (mean = 3.55, SD = 2.03).

To simulate an online US cosmetics store, we created 16 webpages that consist of the manipulated images integrated with the packshot[1] and a brand logo, textual product information, and other product related information such as price, product rating etc. The latter components were all identical across the four conditions of the

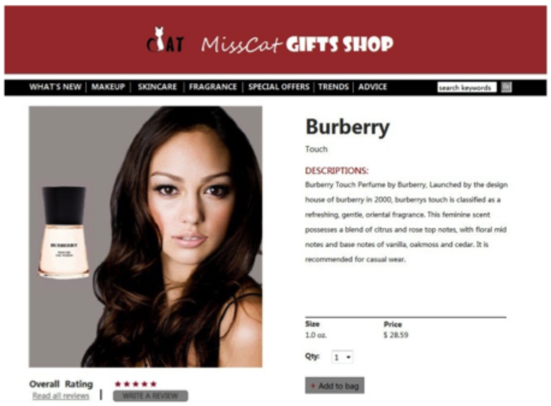
[1] A image of a product, which usually includes the packaging and labeling of the product.



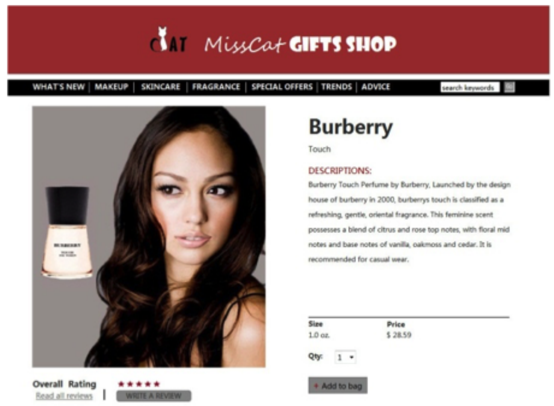
(a) Direct Gaze and Neutral Expression



(b) Averted Gaze and Neutral Expression



(c) Direct Gaze and Happy Expression



(d) Averted Gaze and Happy Expression

Figure 2.1: Examples of the stimuli for experiment 1, with American participants and American fashion models.

study. The webpages were designed with Microsoft Visual Studio 2008. We used two product categories, cosmetics and perfumes, that usually employ human faces on their websites and that the participants are very familiar with. Examples of the stimuli are presented in Figure 2.1.

Participants

We recruited a total of 130 American participants from undergraduate programs at a major U.S. university. They participated in the experiment for course credit or were paid $10 for their participation. Four participants could not be used because they failed to be calibrated in the eye tracking test. Participants were randomly assigned to one of the four treatment conditions. Among the 126 participants, 64 were male and 62 were female, and they had an average age of 21 years. Ninety-three percent of the participants had online shopping experience and 61 percent of the participants shopped online at least once or twice per month.

Experimental procedure

Participants were screened to ensure normal or corrected to normal vision; those wearing hard contact lenses or eye glasses were excluded. The stimuli were presented on a 19 inch monitor with a resolution of 1024×768 pixels. Participants sat individually in a cubicle in front of the monitor at a distance of approximately 60 cm. Eye movements were collected with a Tobii 1750 binocular infrared eye tracker at a frequency of 50 Hz. The eye tracker allows participants to freely move their head in a virtual box of around $30 \times 30 \times 30$ centimeters.

Participants were calibrated, then looked at a fixation cross on a blank page for one second (to get a baseline measurement of pupil size), and were shown an introduction page explaining the task. Then, they were instructed to freely view the web pages: "Imagine that you want to buy some gifts for a female friend and you are freely browsing an online gift store, which sells various kinds of cosmetics and perfumes. A series of product web pages will be presented. Please look through these pages at your own pace." The web pages exactly occupied the entire screen, scrolling was neither needed nor allowed. Each participant viewed 16 pages in a random order, for as long as they desired. After viewing all pages, participants were asked to complete a survey with questions on perception of facial cues, and information on age, gender, and online shopping experience. Completing the experiment took approximately 15 minutes per participant.

Measures

On each of the webpages, we defined five areas of interest: the face of the model, the eyes, the packshot, the brand logo, and the text ([2]). For each AOI, three fixation indicators were computed: fixation counts (reflecting attention capture and engagement [87], and semantic importance [88]), average fixation duration (in msec., reflecting depth and effort of cognitive processing, [89, 90, 88]), and time to first fixation (in msec., reflecting attention selection [91], and conspicuity [88]). The analyses of time to first fixation yielded no significant effects and are not presented in the Results section. We also computed the total dwell time (reflecting overall interest [92, 88]) and the average pupil diameter (reflecting emotional arousal [93], and/or mental workload [88]) for each webpage for each participant. Pupil diameter was averaged across both eyes and computed relative to a baseline measure obtained from a blank page.

Analyses

We used a Bayesian approach to ANOVA [1] to analyze the eye tracking data. Compared to standard ANOVA, Bayesian ANOVA offers the well-known advantages that it provides exact inferences for finite samples of participants, accommodates specific distributional properties of the various eye tracking measures, and includes unobserved differences between participants. Let $i = 1, \ldots, I$ denote participants and $j = 1, \ldots, J$ denote webpages ($I = 126$, $J = 16$). We modeled the various eye tracking measures with the following distributions.

Fixation count data

The fixation count, $y_{i,j}$, indicates attention retention of individual i watching each AOI of j^{th} webpage. It was described with a zero-inflated Poisson distribution, $y_{i,j} \sim ZIP(p_i, \lambda_{i,j})$, introduced by Lambert(1992):

$$f(y_{i,j}, \lambda_{i,j}, p_i) = p_i \text{Pois}(0) + (1 - p_i)\text{Pois}(\lambda_{i,j}), \tag{2.1}$$

where p_i is the probability of a structural zero (indicating attention capture), with $logit(p_i) = \eta_{i,2}$. Nonzero fixation count is assumed to follow a Poisson distribution with fixation rate $\lambda_{i,j} \sim LN(\eta_{i,1}, \sigma_1^2)$, accommodating a log-Normal heterogeneity distribution across participants.

Average fixation duration and total dwell time

The average fixation duration, $f_{i,j}$, which indicates depth of processing, was described by a log-normal distribution:

$$f_{i,j} \sim \text{LogNormal}(\mu_{i,j}, \tau^2), \tag{2.2}$$

where the mean $\mu_{i,j} \sim N(\eta_{i,3}, \sigma_2^2)$ accommodates unobserved differences between participants.

The model for the total dwell time, $g_{i,j}$, as a function of $\eta_{i,4}$ was similar.

Time to first fixation

Time to first fixation, $t_{i,j}$, was right censored data. If the AOI in question is not fixated at all, time to fixation is not observed. Instead we recorded t_{ij} as $\geq$ total dwell time. We modeled it with a Weibull distribution:

$$t_{i,j} = \min(t^*_{i,j}, g_{i,j}) \text{ and } t^*_{i,j} \sim \text{Weibull}(\rho, v_{i,j}), \tag{2.3}$$

where $t^*_{i,j}$ is the true time to first fixation data without censoring. When $t^*_{i,j}$ is smaller than the total dwell time $g_{i,j}$, it is observable and $t_{i,j} = t^*_{i,j}$. Otherwise, it is right censored at $g_{i,j}$. The parameter $v_{i,j} \sim LN(\eta_{i,5}, \sigma_3^2)$ models individual differences.

Average pupil diameter

Finally, average (of two eyes) pupil diameter, $d_{i,j}$, indicating level of arousal, was modeled with a Normal distribution:

$$d_{i,j} \sim N(\kappa_{i,j}, \tau^2), \tag{2.4}$$

where $\kappa_{i,j} \sim N(\eta_{i,5}, \sigma_4^2)$ models individual differences. We imposed non-informative prior $N(0, 100)$ on all σs.

We modeled the between-participant effects of the factors gaze direction (D), facial expression (E), and the D $\times$ E interaction on each of the $p = 1, \ldots, P$ measures described above ($P = 5$), for each of the AOIs. We include the effect of gender (G) as well, because we want to control for differences between men and women and because these effects are also of some interest in their own right. We use the following three-way ANOVA model:

$$\eta_{i,p} = \theta_p^0 + \theta_{p,k_i}^G + \theta_{p,m_i}^D + \theta_{p,n_i}^E + \theta_{p,m_i,n_i}^{DE} \tag{2.5}$$

The parameters θ_{p,k_i}^G represent the effects of gender, where k_i is the gender k (woman/man) of participant i. The parameters θ_{p,m_i}^D represent the effects of gaze direction, where m_i is level m (straight/averted) for participant i. The parameters θ_{p,n_i}^E represent the effects of facial expression, where n_i is level n (neutral/happy) for participant i. Finally, the parameters θ_{p,m_i,n_i}^{DE} represent the interaction effect of gaze direction and facial expression. Because we set $\theta_{p,1}^D = 0$, $\theta_{p,1}^E = 0$, $\theta_{p,m_i,1}^{DE} = 0$ and $\theta_{p,1,n_i}^{DE} = 0$, all effects are relative to the original image with a face with direct gaze and a neutral expression, represented by the parameter θ_p^0.

The models were estimated with JAGS[2], simulating two chains with a burn-in of at least 20,000 and standard methods to check convergence. We report the posterior means and standard deviations of the parameters and compute Bayesian p-values to indicate statistical significance.

Results

The results of the analyses are shown in Table 2.1. Gender had a significant effect on attention to the product and the eye: females tended to have higher fixation counts, and longer fixation durations on the product and the eyes, which may be because the products on the websites were typical female products.

There were no significant effects on attention to the face, for any of the measures. There were significant main effects and significant interaction effects of gaze direction and facial expression on fixation counts and fixation durations on the brand. There was a higher probability that the brand was selected when the model had a neutral expression and averted gaze, compared to a neutral expression and direct gaze. The probability of looking at the brand was also higher for a model with averted gaze and a happy expression, compared to a model with direct gaze and a happy expression. Fixation durations were longer for averted than for direct gaze (neutral expression), and were longer when the model had happy expression than when it had a neutral expression (direct gaze). A model with a happy expression and direct gaze resulted in longer fixation durations on the brand than a model with

[2] http://mcmc-jags.sourceforge.net/

a happy expression and averted gaze.

For the duration of fixations on the text, there was a significant main effect of gaze direction and a significant interaction effect between gaze direction and facial expression. A model with averted gaze resulted in shorter fixation durations on the text than one with direct gaze (neutral expression). Compared to a model with a direct gaze and happy expression, a model with averted gaze and a happy expression resulted in longer fixation durations on the text.

The results for the metrics for the webpage as a whole are shown in Table 2.2. Gender had a significant effect on overall dwell time: compared to female participants, males spent more time looking at each webpage. We also found that males had a larger pupil size overall. This is likely because female models were displayed on the websites.

There was a significant main effect of facial expression on pupil size: participants had a larger pupil size when the model has a happy expression (with direct gaze). There was a significant interaction between gaze direction and expression on pupil size: participants' pupil was wider when the model had a direct gaze and happy expression as compared to one with averted gaze and a happy expression.

Discussion

In this study among American participants we found support for joint effects of facial expression and gaze direction. First, we found that when a model looks at the product with a neutral expression this resulted in a higher probability of

Table 2.1: Results for American (study1) and Chinese (study 2) participants: effects of gaze direction and facial expression on fixation count and fixation duration. The posterior mean and standard errors (in parentheses) are reported. Estimates that have a Bayesian p-value less than 0.05 are in bold

Variable	American			Chinese		
	Fixation Count		Fixation Duration	Fixation Count		Fixation Duration
	Selection	Engagement		Selection	Engagement	
Face						
Intercept	**-13.980 (1.927)**	**1.501 (0.133)**	**5.043 (0.065)**	**-11.382 (3.988)**	**1.186 (0.142)**	**4.935 (0.104)**
Direction (D)	-4.830 (7.859)	-0.234 (0.170)	-0.036 (0.086)	-7.648 (6.944)	-0.064 (0.186)	0.056 (0.136)
Expression (E)	-4.236 (8.072)	-0.073 (0.169)	0.023 (0.085)	**7.218 (3.683)**	0.095 (0.187)	**-0.383(0.136)**
D*E	-2.293 (8.977)	0.311 (0.241)	0.167 (0.121)	-4.104 (8.033)	-0.097 (0.271)	0.326 (0.195)
Gender	-4.784 (7.684)	-0.091 (0.121)	-0.009 (0.061)	**2.450 (2.118)**	-0.102 (0.136)	**-0.332(0.098)**
Eye						
Intercept	**-3.418 (1.087)**	**0.789 (0.174)**	**4.258(0.117)**	**-12.072 (6.072)**	**0.139 (0.244)**	**3.303 (0.154)**
Direction (D)	-6.689 (5.215)	-0.199 (0.225)	-0.180 (0.152)	-4.865 (7.796)	0.081 (0.317)	0.341 (0.202)
Expression (E)	-2.590 (4.030)	-0.027 (0.223)	0.047(0.150)	-2.795 (8.257)	-0.281 (0.319)	**-0.660(0.203)**
D*E	5.274 (7.775)	0.262 (0.319)	0.173(0.213)	-1.420 (9.213)	0.153 (0.460)	0.238 (0.294)
Gender	-1.919 (4.848)	-0.313 (0.160)	**-0.240 (0.107)**	-5.140 (7.546)	-0.184 (0.234)	0.040 (0.146)
Product						
Intercept	**-9.908 (5.218)**	**0.842 (0.136)**	**4.803(0.114)**	**-7.602 (3.936)**	**0.517 (0.236)**	**3.900 (0.152)**
Direction (D)	-7.615 (7.038)	-0.105 (0.175)	0.196(0.145)	**-8.497 (6.288)**	-0.238 (0.310)	-0.152(0.198)
Expression (E)	-2.758 (6.513)	-0.114 (0.176)	-0.082 (0.148)	-5.388 (6.383)	0.011 (0.309)	-0.052(0.201)
D*E	-2.109 (8.735)	0.191 (0.248)	0.038(0.208)	-1.123 (8.938)	0.151 (0.448)	0.280 (0.287)
Gender	0.668(7.106)	**-0.350 (0.127)**	**-0.685 (0.103)**	4.514 (4.646)	-0.317 (0.228)	**-0.508(0.147)**
Brand						
Intercept	**-3.041 (0.326)**	**1.614 (0.117)**	**4.907 (0.073)**	**-6.917 (1.160)**	**1.942 (0.101)**	**5.358 (0.066)**
Direction (D)	**-6.899 (4.049)**	-0.101 (0.151)	**0.263 (0.096)**	1.654 (1.051)	0.036 (0.132)	-0.121(0.086)
Expression (E)	-0.433 (0.416)	0.014 (0.149)	**0.201 (0.094)**	**2.451 (0.976)**	-0.087 (0.132)	**-0.311(0.086)**
D*E	**6.418 (4.076)**	0.080 (0.213)	**-0.394 (0.134)**	**-3.658 (1.554)**	0.149 (0.192)	**0.608 (0.124)**
Gender	0.056 (0.367)	-0.052 (0.107)	0.028 (0.069)	**2.398 (0.667)**	-0.131 (0.097)	**-0.311(0.063)**
Text						
Intercept	**-3.811 (0.346)**	**3.140 (0.133)**	**5.262(0.047)**	**-3.608 (0.417)**	**2.977 (0.154)**	**5.333 (0.050)**
Direction (D)	0.085(0.429)	-0.193 (0.173)	**-0.154 (0.062)**	-0.713 (0.736)	-0.263 (0.203)	**-0.154(0.065)**
Expression (E)	-0.719 (0.516)	0.022 (0.172)	0.065(0.060)	-1.394 (0.869)	0.265 (0.203)	0.078 (0.065)
D*E	-0.286 (0.768)	0.078 (0.245)	**0.189(0.087)**	0.935 (1.337)	0.031 (0.294)	**0.235 (0.094)**
Gender	0.243(0.361)	0.013 (0.122)	0.061(0.043)	-1.151 (0.728)	-0.046 (0.147)	-0.022(0.047)

Table 2.2: Results for American (study 1) and Chinese (study 2) participants: effects of gaze direction and facial expression on dwell time, pupil diameter, blink rate and recall rate. The posterior mean and standard errors (in parentheses) are reported. Estimates that have a Bayesian p-value less than 0.05 are in bold.

	American		**Chinese**		
	Overall Dwell Time	Pupil Diameter	Overall Dwell Time	Pupil Diameter	
Intercept	**9.526 (0.029)**	**3.773 (0.022)**	**9.481 (0.037)**	**3.435 (0.039)**	
Direction (D)	-0.071 (0.038)	-0.008 (0.029)	**-0.135 (0.048)**	0.072 (0.051)	
Expression (E)	-0.007 (0.038)	**0.170 (0.029)**	0.033 (0.048)	0.070 (0.050)	
D*E interaction	-0.013 (0.053)	**-0.087 (0.042)**	0.061 (0.068)	**-0.214 (0.074)**	
Gender	**0.089 (0.027)**	**0.079 (0.021)**	-0.046 (0.035)	**0.154 (0.038)**	
			Blink Probability	Blink Frequency	Recall Rate
		Intercept	**-15.87 (3.938)**	**-1.449 (0.206)**	**0.101 (0.157)**
		Direction (D)	**7.490 (3.054)**	-0.173 (0.271)	**0.483 (0.202)**
		Expression (E)	-7.920 (7.030)	0.042 (0.269)	0.237 (0.199)
		D*E	-5.249 (7.861)	-0.475 (0.391)	-0.033 (0.292)
		Gender	**7.306 (2.954)**	-0.363 (0.201)	-0.028 (0.15)

selecting the brand and deeper processing of the brand. This finding is consistent with our hypotheses and supports the large body of previous research which has demonstrated that gaze cues can effectively orient viewers attention ([39, 40, 41, 42]). It is important to note, however, that the present findings indicate a conceptual rather than perceptual cuing effect ([94, 95]). The brand region in the website is located opposite of the location cued by the model's gaze (Figure 2.1), but in spite of that participants looked more often at the brand when the physical product was cued ([96, 95]). Hutton and Nolte (2011) [17] found a similar conceptual cuing effect to the brand element in an advertisement. Previous findings that to Western participants cues from the eyes are less important may partially explain the absence of a perceptual gaze cueing effect ([64]). The finding that gaze cueing has an effect on fixation durations is new, and suggests that it stimulates the viewer to process the brand deeper when it is cued.

Second, we found that when a model looks at the viewer with a happy expression rather than a neutral one, the resulting average fixation duration on the brand was longer. When a model looks at the product with happy expression (as compared to a neutral expression), the brand received shorter fixation durations. This result is counter to the reinforcement hypothesis of the moderation of gaze cueing by facial expression (for example, Bayliss et al. 2007, [57]). According to Jones et al. (2006) [56] a face looking at the viewer is perceived as more attractive when its' expression is happy than when it is neutral, but the reverse holds for a face with averted gaze. This suggests that in this experiment among American participants, a higher perceived attractiveness of a smiling model that looks at the viewer led to more attention to and to deeper processing of the brand. This supports the contagion theory ([43, 44]).

However, this occurred only for the brand element of the website, while the positive affect caused by a smiling face with direct gaze reduced depth of processing of the product information contained in the text. We speculate that the direction of the emotional contagion effect may depend on the facts/feelings information mode of the description/brand elements, respectively ([97, 98, 99]).

Third, and supporting the above result, we found that while for a neutral expression there is no effect of gaze direction on pupil dilation, for happy expression the pupil is wider if the model looks directly at the participant. This indicates a higher level of excitement if the model smiles and looks directly at the participant.

Previous research has shown that approach-oriented emotions such as happiness are better recognized with direct gaze than with averted gaze ([21]), and that faces that look at the viewer with a happy expression are perceived as more attractive ([56]). Thus, the present results indicate that a model looking at the respondent with a smile elicits positive affect and increases the participants level of arousal, leading to deeper processing of the brand.

In this experiment, we investigated the effects of facial cues for American participants. However, prior research has demonstrated that the perception of emotional expressions (e.g. Jack, Caldara and Schyns 2012; Ozono et al. 2010 [24, 69]) and gaze behavior (Pierson and Bond 1982; Bond and Komai 1976 [71, 70]) are influenced by cultural background. To evaluate differences between Western and Asian cultures, we conducted an experiment in China that replicated the design of experiment 1, but with Asian rather than Caucasian models. Moreover, we added two new measures, brand recall and blink rate. It has been shown that lower blink rates are associated with higher levels of arousal produced by a picture with pleasurable content ([100, 101]), and that blinks occur less frequently when attentional focus is higher ([102]). We added these measures to provide more insights into mental workload, attentional focus and information retention.

2.3 Experiment with Chinese Participants and Chinese Models

In experiment 2, we replicated experiment 1 in China with Chinese participants and Chinese stimuli, using a 2 (facial expression) × 2 (gaze direction) between-subject design.

Stimuli

We obtained from the internet 20 images of female Chinese fashion models shown above the shoulder, facing the camera, and with a neutral facial expression. Adobe Photoshop digital imaging software was used to manipulate the mouth and eyes in these images, respectively, to produce smiling faces and faces that look at the product. In a pretest, 20 participants were asked to rate the images as in experiment 1 and based on the results, 16 images were chosen. Manipulation checks among 80 participants in the main study (see below) revealed a significant difference ($p < 0.001$) in perceived smile between images with a happy expression (mean = 5.33, SD = 1.545) and a neutral expression (mean = 3.66, SD = 1.622). There also was a significant difference ($p < 0.001$) in perceived gaze direction between images with direct gaze (mean = 2.39, SD = 1.829) and those with averted gaze (mean = 5.28, SD = 1.820).

To simulate an online Chinese cosmetics store, we created 16 webpages that consisted of the manipulated images of the faces, integrated with the packshot, brand logo, textual product information, etc.. Text was translated and back-translated into Mandarin Chinese. Examples of the stimuli are presented in Figure 2.2.

(a) Direct Gaze and Neutral Expression

(b) Averted Gaze and Neutral Expression

(c) Direct Gaze and Happy Expression

(d) Averted Gaze and Happy Expression

Figure 2.2: Examples of the stimuli for experiment 2: with Chinese participants and Chinese fashion models.

Participants

Eighty Chinese college students took part in the experiment. The participants were recruited from an online discussion forum affiliated with a university in southern China. Participants were paid \$5 for their participation. Exclusion criteria were similar to those in experiment 1. There were 43 females and 37 males, with an average age of 22.8. All participants had online shopping experience and 74 percent of the participants did online shopping at least once or twice a month. The 80 participants were randomly assigned to four treatment conditions.

Experimental procedure

The procedures in experiment 2 were identical to those in experiment 1. The eye tracking device used in this experiment was the SMI Hi-Speed iView X infrared monocular eye-tracker, with a sampling rate of 500 Hz. The experimental websites were presented on a 19 inch monitor with a resolution of 1024×768 pixels, scrolling was not needed. Participants heads were stabilized using a chin rest at a distance from the monitor of about 60 cm.

The AOIs and eye-tracking measures were the same as in experiment 1. The blink rate was computed as well (measuring mental workload or maintained concentration, Holmqvist et al. 2011, p. 411). The memory test for brand names was added after the survey. Participants were presented with 32 brand names in random order (16 valid brands and 16 invalid brands), and were asked to identify the brand names that they had previously seen. The recall score was calculated by counting the number of correct identifications.

Analysis was done with the Bayesian ANOVA described above. The ZIP model was used for the blink rate (using total view time as an offset), and a Binomial distribution was used for the recall score. The analyses of time to first fixation yielded no significant effects, and are not presented.

Results

The results of the analyses of fixation counts and fixation durations are shown in Table 2.1. First, we found several gender differences: men had a lower probability of selecting the face and the brand, and shorter fixation durations on the face, the product, and the brand. We conjecture that these effects are caused by the product being a "female product", which may lead to a lower level of interest among men.

Gaze direction had significant effects on attention to the product. A model with averted gaze resulted in a higher probability of fixating on the product, indicative of a gaze cueing effect.

Facial expression had a significant effect on attention to the face. Relative to a neutral expression, a smiling expression resulted in a lower probability of fixating on the face and shorter fixation durations. Facial expression also had a significant effect on attention to the eyes: a happy expression resulted in shorter fixation durations.

Facial expression interacted with gaze direction to affect fixation counts and fixation duration. For a face with direct gaze and a happy expression, there was a lower probability that the brand was fixated on and the brand received shorter fixation durations, compared to a neutral expression and direct gaze. However, compared to a face with direct gaze and a happy expression, a face with averted gaze and a happy expression resulted in a higher probability of the brand being fixated on and longer fixation durations on it.

Gaze direction and facial expression also had interactive effects on attention to the text. A face with a neutral expression and averted gaze resulted in shorter fixation durations compared to one with a neutral expression and direct gaze. A face with averted gaze and a happy expression had longer fixation durations on the text, as compared to one with direct gaze and a happy expression.

The results of the analysis of overall dwell time, pupil size and blink rate on the webpage as a whole are shown in Table 2.2. There was gender effect on pupil size and blink rates: men had larger pupil sizes and a lower probability of blinking while looking at the webpages than women. We conjecture that this is caused by a higher level of excitement when looking at the faces of female models.

Gaze direction had a significant effect on overall dwell time: a face with averted gaze resulted in shorter overall gaze duration than one with direct gaze.

There was a significant interaction effect between gaze direction and facial expression on pupil size. A face with direct gaze and a happy expression causes a larger pupil size than one with an averted gaze and a happy expression. In addition, a face with averted gaze resulted in a lower probability of blinking.

Finally, there was a significant effect of gaze direction on recall: a face with averted gaze resulted in a higher recall rate of the brand than one with direct gaze.

Discussion

In experiment 2, we found evidence of the gaze direction cue. As predicted, when a model looks at the product, Chinese participants had a higher probability of fixating on the product, a lower probability of blinking (showing a high level of focus), and a higher level of brand recall. It is likely that Chinese participants orientation of gaze towards the product combined with a higher level of focus led to improved brand recall. These are new results on gaze cueing, and they support and extend the large body of previous research reviewed above. We found no effects on time to first fixation and on fixation counts (indicating engagement), however. Thus, it seems that gaze direction did not affect the speed with which the product was selected for those viewers that look at it, nor the amount of attention paid to it once it was selected. Rather, the gaze cue affected the probabilities with which product and brand information were selected. It thus appears from our results that gaze cues prevent the object towards which gaze is oriented being missed, but do not cause more engagement with it.

With regard to facial expression, we found that a model with a happy expression causes Chinese participants to have a lower probability to look at the face and the brand. It also resulted in shorter fixation durations on the face, the eyes and the brand. These shorter fixation durations led to a shorter overall dwell time on the website as a whole. It seems that a happy expression for Chinese participants causes less deep processing of most elements of the website. To understand this finding, recall that in this study images with a happy expression were created by editing the mouth. Previous research has shown that Asian participants rely strongly on information from the eyes to process smiles ([64]). We speculate that the manipulation of non-Duchenne smiles may have led to lower confidence in the displayed emotion among Chinese participants ([69]), which caused less deep processing of most elements on the website, in particular of the face and the eyes.

We also found support for a joint effect of facial expression and gaze direction. First, we found that for a neutral expression neither the probability of selecting the brand, nor the duration of fixations on the brand, nor the pupil diameter, were different between a model looking at the product or at the viewer. However, a model looking at the product with a smiling face resulted in a higher probability of selecting the brand and longer fixation times on the brand. In this case the fixation duration on the text was also slightly longer. A possible explanation for these findings is that direct gaze facilitates the detection of the non-Duchenne smile ([54]), which leads to lower confidence in the displayed emotion, a weaker contagion effect and less deep processing of brand information. Among Chinese participants, the contagion effect of the smile ([57]) emerges most strongly with an averted gaze.

Comparison of Effects of Facial Cues between American and Chinese Participants

Experiments 1 and 2 revealed attention effects of facial cues that are common across American and Asian cultures. If a models gaze was directed at the product, American and Chinese participants had a higher probability to select the brand

and/or the product. This confirms predictions based on gaze cueing and is in line with previous research. We also found that facial expression moderated the effect of gaze direction on attention to the text, in a similar way for both Chinese and American participants. Specifically, a model with a neutral expression that looks at the product produces shorter fixation durations at text, indicating less deep processing of product information. A model with a happy expression looking at the product, resulted in longer fixation durations on the text. These results are new, and suggest a contagion effect of the facial expression on affect towards the product and brand. This is important given the lack of joint effects in some prior studies ([22]). Finally, there were no effects of face cues on attention engagement and time to first fixation among both Asians and Americans, revealing that the gaze cue mostly prevents the object that was cued (the product or brand) being missed by the viewer, but that there is no effect of gaze direction and facial expression on attention engagement or disengagement in either culture ([49, 22]).

We also found several differences across the two cultures. First, strictly speaking, we found somewhat different gaze cueing effects across different cultures. Chinese participants showed both a perceptual and a conceptual gaze cuing effect whereas American participants showed a conceptual gaze cueing effect only. A possible explanation is that compared to Westerners, Asians focus more on the eyes ([64]), thus are more likely to directly follow the gaze direction of the models.

Second, the effects of facial expression are very different between American and Chinese cultures. When the model on the website smiles, American participants pay more attention to the brand. However, for Chinese participants a smiling face caused a lower probability of selecting the brand, and resulted in less deep processing of most elements on the website. Again, this can be explained because Asian participants preferentially fixate and take cues from the eye region ([62]). This was confirmed by our results: the intercept of the fixation indicator for the models eyes was much more negative in experiment 2 compared to experiment 1 (by almost a factor of four, see Table 2.1), which shows that Chinese compared to Americans have a much higher probability to look at the eyes. To American participants the mouth is the most crucial diagnostic feature in identifying and interpreting smiles ([65, 66, 67]), while Asian participants rely strongly on information from the eyes to detect and interpret smiles ([64, 61]). In our two experiments, images of models with a happy expression were purposely produced by digital editing of the models mouth, producing a non-Duchenne smile ([63]). The second experiment showed that Chinese participants had shorter fixation durations on the eyes of a smiling model, indicating that they processed the information less deeply. Thus, we speculate that American participants paid more attention to the product or brand when the model on the website smiles, because the mouth region of the non-Duchenne smile elicited stronger positive affect, which carried over to the brand. But Chinese participants showed less attention to the brand and processed the information less deeply, because they were less confident about the non-Duchenne smile, taking mostly cues from the eyes. This reversed the contagion effect.

Third, the extent to which facial expression moderates the effect of gaze direction on attention to the brand was different between Americans and Asians. For American participants, a model that looked at the product with a neutral expression drew more attention to the brand than one that looked directly at the viewer, while for Chinese participants there is no difference. For Chinese participants, a model that looked at the product with a happy expression resulted in more attention to the brand. For American participants, taking stronger cues from the mouth, a model looking directly at the viewer with a happy expression elicited a more positive affect towards the model which carried over to the brand. However, for Chinese participants, taking stronger cues from the eyes, it is easier to detect the non-Duchenne smile with direct gaze, which will lead to more uncertainty about the emotion and less attention to the brand.

The comparison between experiment 1 and 2 showed that the most important cultural differences were found were in the effects of facial expression. However, while we used Caucasian models in experiment 1, we used Asian models in experiment 2. According to the research reviewed above, a match between the ethnicity of the model and the viewer might impact the latter's perception and evaluation of the website. Thus, we further investigate whether the differences with Chinese participants that we found are caused by the use of Asian models. Therefore, in experiment 3 we use both Asian and Caucasian models with Chinese participants, to investigate the joint effects of the models facial expression and ethnicity.

2.4 Experiment with Chinese Participants and Various Fashion Models

This experiment was conducted in China with Chinese participants and used a 2×2 between-subject design, with two levels of facial expression (neutral versus happy) and two levels of models ethnicity (Asian versus Caucasian).

Stimuli

We obtained from the first two experiments 32 images of female Western and East Asian fashion models. In a pretest, 13 participants were asked to rate the images on attractiveness, the degree of naturalness, and feelings about the smile. Based on the results, 20 images were chosen as stimuli, which included 10 Caucasian and 10 Asian models. Similar to experiment 2, we created 10 web pages for each condition.

Participants

Sixty Chinese college students took part in the experiment. The participants were recruited from an online discussion forum affiliated with a university in southern China. Participants were paid $5 for their participation. Exclusion criteria were same as those in experiment 2. There were 27 females and 33 males, with an average age of 23.5. All participants had online shopping experience and 95 percent of the participants did online shopping at least once or twice a month. The participants were randomly assigned to the four experimental conditions.

Experimental procedure

The procedure in experiment 3 was identical to that of the first two experiments. The eye-tracking device used was same as that in experiment 2. On each of the web pages, we defined the same five areas of interest as in experiments 1 and 2, and used the same eye-tracking measures. Analyses were done with the Bayesian ANOVA, as described above, with the ethnicity of model and emotional expression as between-subject factors. The analyses of time to first fixation yielded no significant effects, and are not presented.

Results

The results of the analyses of fixation counts and fixation durations are shown in Table 2.3. Gender affected the fixation count on the eye significantly, males having less fixations than females.

There were no significant effects for any of the eye movement measures on the face and the brand. There was a significant main effect of models ethnicity, a marginally significant main effect of models facial expression and a significant interaction between the models ethnicity and facial expression on attention to the product. Compared to an Asian model, a Caucasian model with a neutral expression resulted in a lower probability of fixating on the product. A Caucasian model with a happy expression resulted in longer fixation durations on the product compared to a neutral expression. However, an Asian model with a happy expression resulted in a shorter fixation duration on the product than a neutral expression. There was a significant main effect of the models ethnicity on attention to the text: Caucasian models resulted in a lower probability of fixating on the text.

The results of the analyses of overall gaze time, pupil size and blink rate on the webpage as a whole are shown in Table 2.4. The models ethnicity had a significant effect on blink rate and pupil size. Asian models resulted in a lower rate of blinking and a larger pupil size. Facial expression also had a significant effect on pupil size. Compared to a model with a neutral expression, a model with a happy expression caused a larger pupil size.

Discussion

In experiment 3, we found that among Chinese participants, relative to a Caucasian model an Asian model produced a higher probability of fixating on the product and text, elicited a lower rate of blinking and a larger pupil size, and resulted in a longer time browsing the web page. The results seem to indicate that Chinese participants considered content from a website with an Asian model (in-group) was personally relevant and targeted to them, which led to more attention to product information and the webpage as a whole. These findings are consistent with our prediction and support previous research that has demonstrated positive racial in-group effects ([82, 77, 79, 25, 78, 80, 81]).

Further, we also found that the models ethnicity moderated the effect of her facial expression. An Asian model with a happy expression resulted in shorter fixation durations on the product, as compared to a neutral expression, indicating less deep processing of information. This is consistent with the findings in experiment 2. We speculated that a non-Duchenne smile may have led to lower confidence in the displayed emotion among Chinese participants ([69]). Brown, Bradley and Lang (2006) [86] found that people had stronger affective responses when viewing pictures of ethnic in-group as compared to ethnic out-group members. Chinese may thus have had lower confidence in the displayed emotion when viewing web pages with Chinese models with a non-Duchenne smile (ethnic in-group) than with Western models (ethnic out-group). This would have led to less deep processing of product information on websites that contained Chinese models with a happy expression.

2.5 General Discussion

The findings reported in this paper make several theoretical contributions. First, this research empirically explores the joint effects of facial expression and gaze direction of online models on a wide range of indicators of consumers attention in an e-commerce context. Although many studies in psychology have investigated the effects of these two facial cues ([46, 43, 14, 15, 47]), there has been little research effort directed at systematically exploring their joint impact in naturalistic contexts, and it remained unclear whether joint effects existed. First, the present studies added to previous research on gaze cueing by showing that the cuing effect of gaze direction appears to operate via attentional selection, rather than through engagement. It appears from the results that the gaze cue mostly primes initial attention to objects in the visual field that might otherwise be missed. Further, the studies revealed evidence of interactive effects between facial expression and gaze direction of online models, which is important given the discrepancy in findings between prior studies ([21, 36, 22]). The present findings support the contagion theory ([43, 44]), and show that positive affect from the happy expression when a model looked at the viewer carried over to the product or brand. This caused more attention to these elements of the website.

Second, this study established similarities and differences in the effects of facial cues among American and Chinese consumers ([23, 62, 24, 64]). Among both American and Chinese participants, a model with direct gaze and a happy expression causes a larger pupil size than one with an averted gaze and a happy expression, indicating that this combination of facial cues elicits a higher level of excitement. Among both cultures, there was evidence of both perceptual and conceptual gaze cueing, where gaze directed at the product on the website promoted attention to the product and the brand. Facial expression moderated the effect of gaze direction on attention to the text containing product information, in a similar way for both Chinese and American participants. There were also differences between the two cultures. For American participants, a model that looked at the viewer with a happy expression drew more attention to the brand, while for Chinese participants a model that looked at the product with a happy expression resulted in more attention to the brand. These differences are caused by cultural difference in using the eyes and mouth as cues to recognize and interpret smiles in Asian and Western cultures, respectively. Further, the racial match between a model on a website and the viewer exacerbated the attention effects of facial expression (see Brown, Bradley and Lang 2006 [86]).

Third, compared to prior eye-tracking research ([9, 17, 7]) we examined the effects of facial cues on attention to a wider range of regions of interest, and used a wider range of indicators of visual attention. This provides more comprehensive insights into the simultaneous effects of facial cues on multiple processes, including attention selection and engagement, depth of processing, attentional focus, affect, and memory. Importantly, among Chinese participants, our research first revealed memory effects of gaze cueing.

The findings of this study have important practical implications for the design of e-commerce websites. Many of these websites display models with the intention to attract attention and direct it to the product, especially for categories such as apparel, fragrances, accessories, cosmetics and other beauty products. A small content analysis of about fifty web sites shows that compared to Western websites for beauty products, Asian websites employ human images much more frequently, and most models are Asian. On Asian websites for cosmetics, most models have averted gaze, but dont direct their gaze at the product in question. Taking apparel as another example, models with gaze directed at the viewer are more common on Western than on Asian websites. However, the gaze cueing effects demonstrated in this research suggest that a models gaze, when directed at the product, can be used to orient viewers attention to the product or brand in both cultures. In addition, the practice of using smiling faces is different across different websites. Take apparel as an example: almost three quarters of the models on eBay's website shows a neutral expression, whereas over three quarters of the models on Macys website displays a happy expression. Hence, there does not appear to be a consensus on the best choice of models facial expressions. The present results show, however, that for American participants, a smiling model causes more excitement and draws attention to the brand. Moreover, this study suggests that online retailers need to consider interactive effects between these facial cues. For example, online retailers can elicit consumers interest in the content and lead them to process the textual product information deeper, by using models with a happy expression looking at the product.

Further, the current research suggests that online retailers need to consider the impact of facial cues in relation to culture. It is known that websites adapted to the local culture have the potential to make consumers remain longer at the site ([9]). Adapting websites to the preferences of users with different cultural background, therefore, may greatly increase the effectiveness of the website in question, especially given the escalating numbers of global online shoppers. With the development of technologies such as website morphing, automatically changing the characteristics of a website to match culture and demographics has become feasible and practical ([103, 104]). Online retailers can customize their websites by changing the models ethnicity and facial cues to match consumers cultural background when they interact with the website. For example, if the IP address identifies a viewer as Asian, the website can automatically present Asian models with a smiling face looking at the product to elicit more positive affect and more attention to the brand. Gaze cueing can have even more pronounced effects if rather than a static cue, the cue is dynamic by moving the models eyes in the direction of the product.

Table 2.3: Results for Chinese participants (study 3): effects of ethnicity (race) and facial expression on fixation count and fixation duration. The posterior mean and standard errors (in parentheses) are reported. Estimates that have a Bayesian p-value less than 0.05 are in bold.

Variable	Fixation Count		Fixation Duration
	Selection	Engagement	
Face			
Intercept	**-13.555 (5.861)**	**1.353 (0.169)**	**4.815 (0.304)**
Race (R)	-2.532 (7.870)	-0.02 (0.203)	0.099 (0.367)
Expression (E)	-3.281 (7.625)	0.152 (0.205)	0.281 (0.368)
R*E Interaction	0.705 (9.167)	-0.051 (0.288)	-0.173 (0.519)
Gender	-3.465 (7.328)	-0.079 (0.146)	-0.072 (0.265)
Eye			
Intercept	**-11.081 (5.88)**	-0.153 (0.323)	**2.949 (0.492)**
Race (R)	-5.075 (7.922)	0.048 (0.393)	-0.007 (0.593)
Expression (E)	-0.692 (8.540)	0.258 (0.393)	0.262 (0.596)
R*E Interaction	-2.686 (8.853)	-0.080 (0.550)	0.229 (0.838)
Gender	-4.942 (7.501)	**-0.588 (0.279)**	-0.777 (0.430)
Product			
Intercept	**-10.066 (4.202)**	0.412 (0.312)	**4.126 (0.524)**
Race (R)	**6.315 (3.503)**	0.052 (0.378)	-0.879 (0.632)
Expression (E)	-4.803 (7.194)	-0.564 (0.381)	-1.094 (0.636)
R*E Interaction	-2.402 (7.705)	0.674 (0.534)	**2.058 (0.892)**
Gender	2.643 (3.232)	-0.019 (0.27)	-0.084 (0.459)
Brand			
Intercept	**-9.508 (4.334)**	**2.212 (0.202)**	**5.532 (0.28)**
Race (R)	-3.095 (7.214)	-0.003 (0.243)	-0.077 (0.339)
Expression (E)	4.058 (4.341)	-0.351 (0.246)	-0.381 (0.340)
R*E Interaction	-4.143 (7.905)	0.084 (0.346)	-0.061 (0.480)
Gender	1.287 (3.234)	-0.044 (0.180)	-0.413 (0.244)
Text			
Intercept	**-10.097 (3.919)**	**2.950 (0.257)**	**5.595 (0.276)**
Race (R)	**7.154 (3.92)**	0.015 (0.313)	-0.429 (0.333)
Expression (E)	-8.061 (7.223)	-0.159 (0.318)	-0.409 (0.334)
R*E Interaction	-5.546 (7.968)	0.231 (0.447)	0.756 (0.470)
Gender	-0.386 (0.904)	-0.096 (0.224)	-0.226 (0.240)

Table 2.4: Results for Chinese participants (study 3): effects of ethnicity (race) and facial expression on dwell time, pupil diameter, and blink rate. The posterior mean and standard errors (in parentheses) are reported. Estimates that have a Bayesian p-value less than 0.05 are in bold.

Variable	Blink Rate		Dwell Time	Pupil Diameter
	Probability	Frequency		
Intercept	**-10.784 (5.436)**	**-2.313 (0.353)**	**9.571 (0.129)**	**3.718 (0.074)**
Race (R)	-5.39 (6.964)	**0.961 (0.423)**	-0.284 (0.155)	**-0.225 (0.086)**
Expression (E)	-6.081 (6.774)	0.377 (0.424)	-0.032 (0.156)	**0.180 (0.086)**
R*E Interaction	1.169 (9.508)	-0.507 (0.593)	0.299 (0.22)	-0.060 (0.121)
Gender	1.574 (7.006)	0.167 (0.303)	0.166 (0.112)	0.023 (0.064)

Chapter 3: Impact of Emotions on Watching Intention in Comedy Movie Trailers

3.1 Introduction

A movie trailer is a mini-movie that usually contains its own narrative and its own score. It usually contains highly condensed series of selected footage from the movie it is promoting ([105, 106]). By selecting and showing actual scenes from the movie, it signals the quality of the movie and purports to attract viewers. These fragmentary scene cuts allow the audiences to build expectations and imagine the actual film plot ([107, 108]). Voice-over narration is used in movie trailer to explain the story therefore helps audiences to understand the movie better. Trailer music, also called signature music, is often not from the movies soundtrack itself since trailers are usually released long before the score of the movie is composed. The main purpose of trailer music is to provoke or support an emotional reaction to the trailer from the audience. Trailer music has been claimed to be under-appreciated, as only one fourth of the trailers produce their own original score. The others, to save money and time, commonly use the score of other movies, or use classical music, pop songs, or library music ([109]).

Comedy is the leading genre in term of total gross revenue from 1995 to 2014. In this time period, a total of 2,032 comedies have been produced, which is the second most productive genre, after drama, which has about 3,744 movies. The total box office gross of comedy is around $38 billion and the average gross is about $20 million per movie. Comedy has a market share of over 20%, larger than the market share of any other genre. Clearly, the main goal of comedy movies is to amuse and elicit laughter from the audience. There are many comedy subgenres, for example, slapstick, spoof film, black comedy, and romantic comedy, as well as many hybrid genres, such as action comedy, horror comedy, fantasy comedy, and sci-fi comedy.

The movie industry and box office success has seen ample research in marketing. Studies conducted by Litman [1983] [110], and later by Litman and Kohl [1989] [111] have confirmed the relationships between box office success and determinants such as time of release, distributor, movie genre, production costs, and Academy Awards. Sharda and Delen [2006] [112] showed that the success of a movie is determined by the number of screens (on which the movie was shown during the initial launch), special technical effects and superstars. Reviews from critics have also been shown to play a role in influencing the box office ([113]. Conflicting results have been reported for the effect of film ratings. Ravid [1999] [114] found that MPAA ratings of G and PG movies have a positive effect on movie success, while studies by Litman [1983], Austin [1984], Austin and Gordon [1987] [110, 115, 116] showed no significant correlations. Levene [1992] [117] showed that theatre trailers and television advertising were the two most important determinants of box office performance. Faber and OGuinn [1984] [26] confirmed the effect of movie previews and movie excerpts, rather than film advertising, word-of-mouth and critics review, on movie going behavior. Eliashberg et al. [2007] [118] demonstrated further that the scripts in spoilers could be used to forecast a movies return on investment (ROI).

Movies elicit a wide range of strong emotions, including happiness, sadness, surprise, anger, disgust, and fear ([119, 120]). They provide the audience the concentrated experience of these powerful emotions that are not often encountered in day-to-day occasions. Researchers in marketing have studied emotions and shown their impact on decision-making (for example, Isen and Patrick [1983], Lerner et al. [2004], Lerner and Keltner [2000] [121], [122], [123]). Comedies provide the pleasurable experience by using comic devices in the movie, such as jokes, parody, exaggerated behaviors and so on. Like the movies they promote, comedy trailers are targeted at inducing positive pleasurable emotions such as happiness and surprise from their audience. Happiness, as a positive pleasant emotion usually comes from encountering unexpected positive events ([124]). Surprise is a startle-response when an unexpected event occurs, and it can have a negative or positive valence, and thus can enhance or reduce attraction ([125]). However, comedies may also contain offensive and unpleasant humors, for example, slapstick, sarcasm, prejudices and bathroom humor. They can sometimes induce emotions such as disgust. In a study conducted by McGraw and Warren [2010] [126], they showed that benign moral violations tend to elicit mixed emotions of amusement and disgust.

Measuring emotions has been a complex problem ([127]). The most widely used method to measure the subjective emotion experiences is self-report ([128]).

Researchers also have used other objective methods to measure the peripheral physiological autonomic nervous system (ANS) responses of emotions, and have used Electroencephalography (EEG), functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) techniques to measure central physiological (CNS) responses ([129, 130, 131, 132]). Whole bodily posture measures, facial expression and other emotional expression including vocal characteristics are most common ways in psychology to study the complex behavioral responses of emotions ([133, 134]).

The connection between emotions and facial expressions in humans and animals was first studied by Darwin, in 1872. A century later, in 1978, Ekman and Friesen [135] continued the study in facial expression by developing the Facial Action Coding System (FACS) to systematically categorize emotions by coding instant facial muscular changes ([135]). Facial behaviors can be measured by trained coders, who use images of the face and identify human emotions by detecting facial muscle movement, following the FACS Manual. More accurately, facial Electromyography (EMG) is used to detect the electrical activity produced by facial muscles ([136]). Moreover, nowadays specialized automated detection software makes it possible to process real-time video data at a rate of four times per second ([1]).

In previous studies, Teixeira et al. [2012] [4] examined the effects of joy and surprise on concentrating attentions and retaining viewers in online video ads. In a lab experiment, they used eye-trackers to monitor the concentration of attention during exposure to video ads. Facial expression footage was collected and then analyzed by the emotion detection software. Zapping behavior was recorded as a measurement of viewer retention. They found that the surprise level has a stronger effect on attention concentration than joy. The changes in joy impacted viewer retention more than surprise, however. The level of joy improved retention more over time, while the level of surprise improved retention less over time. Teixeira et al. [2014] [5] conducted a large-scale field study on TV advertisements by using a web-based face tracking system. Their study confirmed the linear relationship between entertainment level and viewing interest. It also revealed an inverted U-shape relationship between entertainment level and purchase intent, which shows that too much entertainment could potentially hurt purchase intent. Furthermore, the results also suggested that when entertainment comes after the brand this increases purchase intent, while when the entertainment comes before the brand it does not.

Our study is the first web-based face tracking field study on movie trailers. Our main goal is to investigate happiness induced by movie trailers and its downstream effects on the intentions to watch the movie. To that end, we use automated video analysis data collected with nViso, a marketing research company specialized in web-based emotion recognition. nViso captures facial movements passively via a webcam in participants own home or workplace, while they are watching movie trailers. These moment-to-moment emotion data were collected by cloud computing using an automated facial expression recognition system. This experiment provides data on participants natural reaction to movie trailers, in a setting where they would normally watch them. While multiple emotions measures were collected, in this study we focus on happiness, surprise and disgust, as the key intended emotions of trailers for comedy film. Pleasurable emotions elicited during exposure to multiple long comedy trailers are less intrusive compared to negative emotions, and thus these trailers can be used in studies that last longer periods of time, in the present case up to 45 minutes.

3.2 Experiment Setup and Variable Description

During the experiment, participants were asked to view a webpage that contains a series of 12 comedy movie trailers. Their facial expressions during the exposure were recorded remotely through the webcams. Participants were asked to answer questions about their evaluations of all the trailers and movies as well as their intentions of watching movies.

3.2.1 Participants and Stimuli

A total of 122 paid participants were recruited online by two major northeastern Universities. This group of participants had a mean age of 24 and range from 18 to 68, with 28% males. Each participant received $5 in the form of an Amazon Gift Card if they completed the experiment. To make the experiment incentive-compatible, the participants also had a 1 in 10 chance of winning a free online movie from the movies that he/she chose in the final questions. The eligible participants must have access to a personal computer with a webcam and high-speed Internet connection, have perfect or near-perfect vision without glasses, or use contact lenses. No eyeglasses were allowed. In addition, male participants with large amount of facial hair (full mustache or beard) may not be eligible to participate.

Stimuli were comedy trailers selected from a pool of 100 comedy trailers through a balanced incomplete block design. One randomly selected 2012 comedy trailer was used as a control stimulus, which was shown to all participants at the beginning of the experiment. The webpage of the experiment was designed by nViso. All comedy movie trailers were taken from public access video channels. The movie trailers included green band and red band versions, according to Motion Picture Association of America (MPAA) standards. A green band trailer shows an all-green graphic at the head of the trailer. It indicates that MPAA had approved the trailer for appropriate audience. A red band trailer is for R-rated, NC-17-rated or unrated movies, viewing by restricted or mature audiences, and may not be appropriate for children. Different versions of trailers for the same movie were included, which requires the participants to be at least 18 years of age. Overall, 100 trailers for 50 comedies were used in the study (49 of them were from 2012 movies, 1 from 2011 to make the number 50). Each participant was only exposed to one version of the trailer for the same movie. Thirteen comedy subgenres were selected in the study, including 9 drama comedies, 8 animation comedies, 7 action comedies, 7 romantic come- dies, 4 horror comedies, 4 indie comedies, 4 parodies, 2 black comedies, one for each of political comedy, sci-fi comedy, slapstick, sports comedy and late night comedy. The sets of comedy movie trailers provided enough opportunities for the viewers to be exposed to different levels and types of sub-genres of humor through the trailers.

3.2.2 Procedure

The participants were told that the purpose of the study is to discover features of good comedy trailers. All participants must have a computer that has a functioning webcam. To guarantee the quality of the facial reaction video data, the participants needed to be in a well-lighted space to finish the experiment. Ideally, the participants should be at least 60cm (2ft) away from their webcams. All other applications on their computer such as email, instant message, other browser windows, etc., needed to be closed during the experiment. Further, the participants should refrain from eating, chewing, drinking or talking during the recording process. The participants were told that the compensation would be contingent upon them closely following these instructions.

The participants were asked for consent to participating in the experiment and being recorded via webcam, completing it in one sitting and allowing researchers to use the webcam video for research purposes. After the participants signed the online informed consent form, they clicked on the link to get to the webpage containing the comedy trailers. Each individual was shown a series of 12 trailers randomly selected from a list produced by balanced incomplete block design. This design tried to minimize the spillover effects by randomizing the order of the trailers shown to each participant. The length of each trailer was between 1 to 3 minutes. After each trailer, the participants were asked 5 questions about their previous exposure and their evaluation of the trailer and the movie, and also their intent of watching the movie. After all trailers were shown to the participants, they were asked to answer questions about their age, gender, income, their preference of movie and comedy and their usual movie going behavior. The questions are shown in Table 3.1. By the end of the experiment, the participants were offered a raffle in which they had a one in ten chance to win a free DVD. The participants would choose one or more movies from any of movies they had just watched in the experiment. If they won, one movie was selected from the choices they made. An Amazon gift card for the amount needed to purchase that movie was emailed to the participants. The whole experiment took up to 45 minutes.

3.2.3 Data Description

The videos with facial reaction recorded were analyzed by nViso, which provides cloud computing and real-time operations to measure consumers emotion reactions in their online experimental environment. Machine learning algorithms were used to classify the second-by-second facial muscle movement into seven emotions based on Ekmans Facial Action Coding System ([6]). An emotional profile for each participant, containing the measures of happiness, surprise, sadness, fear, anger, disgust and neutral, was created. The output for each emotion at each time point is a probability indicating the emotion intensity.

We collected two sets of data, one for the online questionnaire responses (see 3.1) and one for the emotion measurements. Some of the collected data did not qualify due to participants relatively poor compliance with instructions (e.g. dark lighting, sitting too close to the camera, interruption during the experiment and

Table 3.1: Questionnaire Measurements

Variable	Question	Measurements
Questionnaire 1		
SeenTrailer	Have you seen this movie trailer before?	Binary, “no” or “yes”
SeenMovie	Have you seen this movie before?	Binary, “no” or “yes”
RateTrailer	How much do you like this movie trailer?	7-point scale, anchored by “not at all” to “extremely”
RateMovie	How would you rate the movie based on this trailer?	7-point scale, anchored by “not at all” to “extremely”
WatchIntention	Would you like to watch this movie?	7-point scale, anchored by “not at all” to “extremely”
Questionnaire 2		
Age	Age	Integer
Gender	What is your gender?	Binary, “female” or “male”
GenrePreference	What genres of movie do you enjoy?	9 categories of movie genres
SubgenrePreference	What subgenres of comedies do you enjoy?	11 categories of comedy subgenres
Inference	What influences you the most when you are trying to figure out whether or not to see a movie?	5 categories (“trailer”, “stars”, “director”, “review” and “family/friend”)
DVD	You could win a DVD of one of the movies you just watched. Check all applied.	12 categories of all movies watched

failure to complete the whole experiment). There were 122 unique respondents in the online questionnaire data and 104 unique respondents in emotion data sets. There were 90 common respondents in the merged data, which means that only 90 of the participants completed the entire questionnaire and had a qualified emotion profile. After merging with control data from the calibration trailer, there were 86 respondent IDs left. One of respondent was eliminated from the analysis due to abnormal emotion measurements in the calibration trailer (the happiness curve is flat). Eventually, only 85 respondents were included in the final data set, which was about 70% of the initially recruited participants.

Dependent Variables Four dependent variables were measured in this study:

- The first two dependent variables, the evaluation of trailer and movie (`LikeTrailer`, RateMovie) are 7-point scale measurements of overall liking of the trailer and a rating of the movie. These two variables are highly correlated at a level of 0.899, confirming the effect of trailer as a signal of the quality of the movie.

- Watching intention (WatchMovie) is also measured on a 7-point scale, indicating how much participants would like to watch the movie (in the theater or on DVD and Blue-Ray) after they have been exposed to the trailer. The watching intention is also highly correlated with the evaluation of trailer and movie (0.870 between WatchMovie and LikeTrailer, 0.892 between WatchMovie and RateMovie). The quality of the movie is the main factor in determining whether participants want to see the movie.

- The final dependent variable is the DVD choice (DVD) that participants entered during the raffle. The participants either choose to enter the raffle with a chance to win the DVD or not. The correlation between the DVD choice and the evaluation and watching intentions were 0.441,0.452, and 0.484 for LikeTrailer, RateMovie and WatchMovie, respectively.

Due to the high correlations between the four dependent variables, currently we only focus on watching intention in our study. The rest of the variables can be analyzed in the same manner.

Movie Trailer Content The movie trailer content was analyzed in terms of the image data and audio data. This yielded the following variables.

- Scene cut: One of the generic features of movie trailer is montage. A trailer usually contains multiple scene cuts from the movie. Figure 3.1 illustrates the scenes captured every second by iMovie from a sample trailer: "Men in Black 3". In this study, scene cuts in the movie trailers are detected automatically using the Scene Detector[1], which is an advanced software utility that detects the scene boundaries based solely on the frame image data. The total number of scenes and the average length of scenes across each trailer are calculated. The location of the longest scene relative to overall number of scenes in the trailer is recorded.

- Audio Volume: We included two types of volume data: (1) Total Volume Data: The MP4 video files are first converted to MP3 audio files. Then the amplitude data are extracted every millisecond using sound processing software SoX[2]. The averages of the absolute values are calculated within each second. (2) Music Volume Data: By removing vocals (as well as drums and bass) utilizing SoX, music throughout each trailer is separated from the audio files, and its volume is calculated as described above. The two types of volume data from the video is relative volume in the trailer video files, which are the same across people, not the self-controlled volume that the participant is listening. For both of the total volume data and music volume data, we calculated the average volume, the slope of volume across the time, the average volume in the start and end scene, and also the scene with highest total and music volume, respectively. Figure 3.2 shows an example of total volume and music volume

[1] http://www.scene-detector.com

[2] http://sox.sourceforge.net

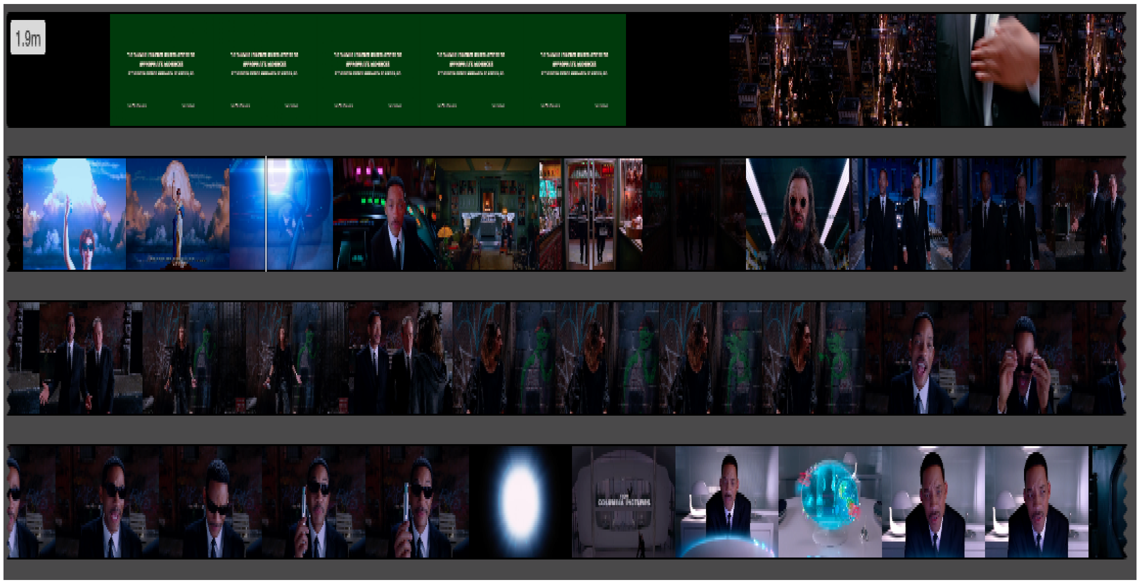


Figure 3.1: Scene cuts from sample trailer (Men in Black 3)

from the same movie trailer with vertical dash line indicating the scene cuts identified in the last step.

Emotion Measures Emotion intensities for each participant were measured on a second by second basis. The levels of fear and anger are relatively low during exposure to comedy trailers. Also sadness is highly negatively correlated with happiness level. Therefore we only consider three emotions: happiness, surprise and disgust in our analysis. Figure 3.3 shows an example of emotion curves containing these three emotions for each participant watching the sample trailer: Men in Black 3. Among these three emotions, the focus of this study is the happiness level due to the high levels cross all trailers. Aggregated emotion measures across the trailer are calculated as follows 3.4:

- Average: We use area under curve (AUC), which is the integrated area under

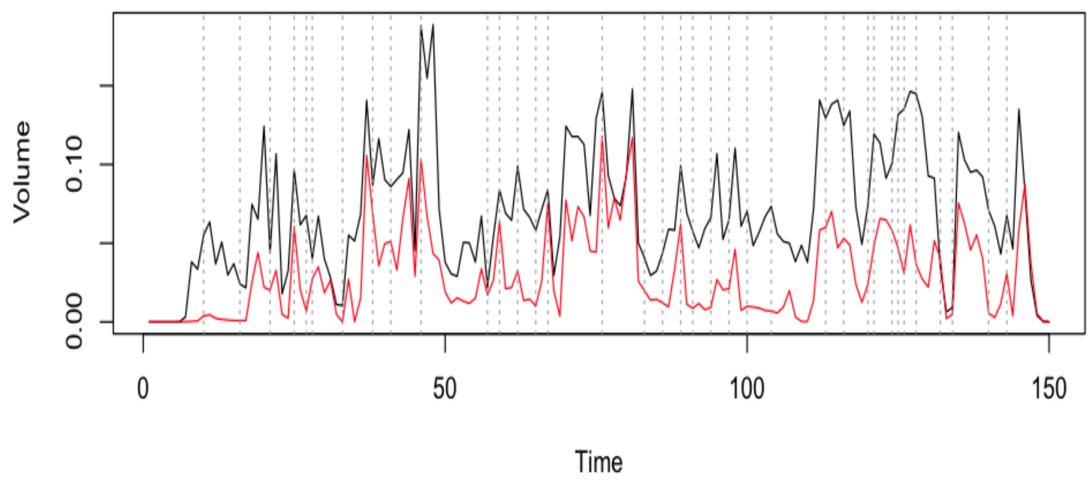


Figure 3.2: Total and music volume. Black line indicates the total volume. Red line is the music volume with vocal removed. Vertical dash lines indicate the scene cuts.

the emotion curve for each individual i watching trailer j to calculate the average emotion.

- Slope: A linear straight line is fit to each emotion curve. The slope of each emotion over time indicates the rate of emotion progress.

- Peak: The scene with the highest average emotion levels is considered to be the peak scene. The average emotion within this scene and the location of the scene relative to the overall number of scenes in the trailer are recorded. The number of spikes and duration of spikes in the emotion curve that exceeds 75% of its peak value is calculated.

- End: Often times, the last scene of the trailer contains montage of strong emotional cuts from the movies. Therefore we also include the average emotion during exposure to the last scene.

- Start: The average emotions during exposure to the first scene in the trailer are also included in our analysis.

Control Variables Previous exposure to the trailer and movie might lead to different emotional responses during the experiment. Personal preference for the movie could affect both emotional response and watching intention after exposure to the trailer. Previous research has demonstrated that men and women differed in movie attendance ([137]) and genre preferences ([138]). Age has also been reported to have an inverse relationship with movie attendance ([137]). Therefore, we use the following control variables:

- `SeenTrailer`: Previous exposure to the trailer (0 = not seen, 1 = seen).
- `SeenMovie`: Previous exposure to the movie (0 = not seen, 1 = seen).
- `GenrePreference`: Preference of movie genres (0 = not preferred, 1 = preferred).
- `SubgenrePreference`: Preference of comedy subgenres (0 = not preferred, 1 = preferred)).
- `InfluenceTrailer`: Whether the participant consider movie trailer to be the most important factor in determining to watch a movie. (0 = no, 1 = yet).
- `Gender`: (1 = male, 2 = female).
- `Age`: Individuals age in years.

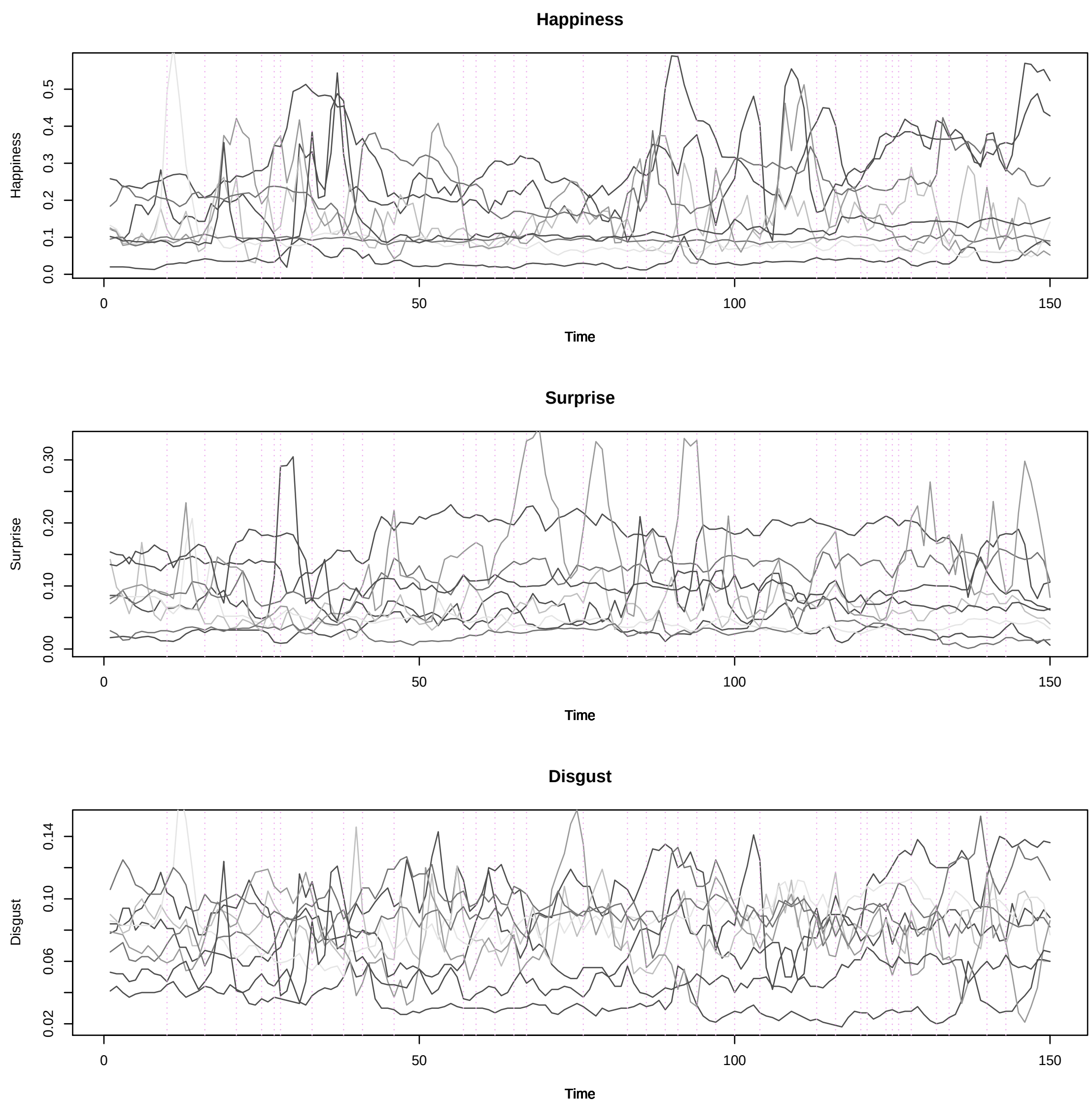


Figure 3.3: Emotion curve for sample trailer (Happiness, surprise and disgust). Darkness of the emotion curve indicates the watching intention for each individual. As watching intention increases, the emotion curve is darker. Vertical dash lines indicate scene cuts.

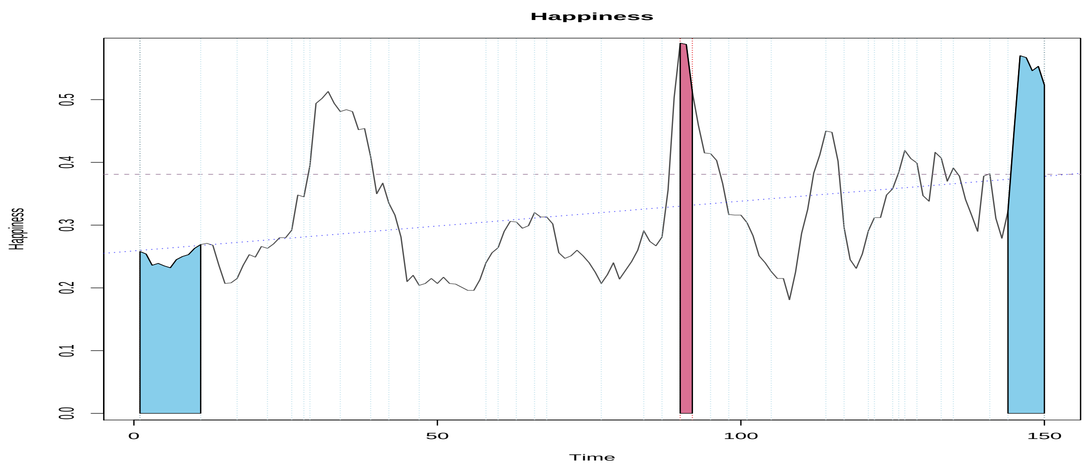


Figure 3.4: Happiness profile for one individual. Vertical dash lines indicate scene cuts. The red shaded area is the region of happiness peak scene. End and start scene are shaded in blue. The red horizontal line indicates 75% of the peak value. Regions above the line are considered to be the spike regions. The blue line is a linear fit of emotion curve, containing information of the intercept and slope.

- `Income`: Personal income numeric categories.

We also obtain control variable data for movies from the "Internet Movie Database" (IMDb) and "The numbers", which are online databases of films, TVs and video games. Specifically, we include release time (whether the movie is released during summer or Christmas holiday season), director productivity (numeric measures of total number of movies directed by the same director) and star popularity (numeric measures of total number of movies played by the same main actor/actress).

3.3 Model

The statistical methodology used in this paper is a joint model that consists of a Bayesian longitudinal model for emotion measures, and a generalized linear mixed model describing the discrete intention to watch the movie, accounting for individual heterogeneity. We first applied a Bayesian variable selection (BVS) method to select the key predictors of watching intention. Then we conducted model selection to select the random effects that link the emotion model and watching intention model. Finally, we compared the model with separate model components with the joint model that includes the selected predictor variables.

3.3.1 Model Components for Emotion and Intention

3.3.1.1 Longitudinal Model for Emotion Data

We used a generalized linear mixed effects model for the continuous emotion measures. The logit function of happiness probability measurement for individual i watching trailer j at time t is denoted as h_{ijt}, which is modeled as:

$$h_{ijt} = \theta_{ijt} + \epsilon_{ijt}, \tag{3.1}$$

where the error terms ϵ_{ijt} are assumed to be independently normally distributed. The underlying emotion trajectory θ_{ijt} is expressed as

$$\theta_{ijt} = \mathbf{W}_{1i}(t) + \mathbf{X}_{ij}^{1}\beta + \zeta_1 S_{jt} + \zeta_2 \mathrm{V}_{jt} + \zeta_3 \mathrm{M}_{jt}, \tag{3.2}$$

where $\mathbf{W}_{1i}(t) = \mathbf{c}_{1i}^{T}(t)\mathbf{U}_i$ are subject-specific random effects, and may include random effects for both the intercept and the slope over time. The matrix X_{ij}^{1} includes individual and trailer specific information described above, including an individuals preference of comedy subgenres and previous exposure to the movie or trailer. The term S_{jt} represents the index of the scene at time t. The term V_{jt} represents the volume of movie trailer j at time t, and M_{jt} represents the volume of only music after vocals are removed from the video. Let $\mathbf{H}_{ij} = (h_{ij1}, h_{ij2}, \ldots, h_{ijT})$ and $\mathbf{U}_i = (u_{i1}, ..., u_{iq})$. The likelihood for the emotion model is:

$$L_{\mathbf{H}_{ij}}(\mathbf{H}_{ij}|\mathbf{U}_{ij}) = \prod_{t=1}^{T} p(h_{ijt}|\mathbf{U}_i), \tag{3.3}$$

where $p(h_{ijt}|\mathbf{U}_i) \sim N(\theta_{ijt}, \sigma_h^2)$.

3.3.1.2 Model for Watching Intentions

An ordered logit model is used for the watching intentions, which are measured on a scale from 1 to 7, with 7 being the highest intention to watch the movie. Let y_{ij} be individual is intention of watching the movie j. A censored latent variable y^*_{ij} is introduced and modeled as follows:

$$y^*_{ij} = X^2_{ij}\alpha + \mathbf{W}_{2i}, \tag{3.4}$$

$$y_{ij} = \begin{cases} 1 & y^*_{ij} < \tau_1 \\ d & \tau_d < y^*_{ij} < \tau_{d+1},\ d = 1, ..., D-1 \\ D & y^*_{ij} > \tau_D, \end{cases} \tag{3.5}$$

where $\mathbf{W}_{2i} = d^T_{1i}\mathbf{U}_i$ contains the subject-specific effects that are similar to the form of $\mathbf{W}_{1i}(t)$. The matrix $\mathbf{X}^2_{ij}$ consists of fourteen predictor variables selected from the Bayesian variable selection procedure, including aggregated emotion measures, and the volume and music data extracted from the movie trailers. The threshold parameters satisfy the order constraint: $\tau_1 < \tau_2 < ... < \tau_D$, where $D = 7$.

The probabilities of each intention category are modeled as:

$$Pr[y_{ij} = 1] = Pr(y^*_{ij} < \tau_1) = Pr(\mathbf{X}^2_{ij}\alpha + \mathbf{W}_{2i} < \tau_1) = \Phi\big(\frac{\tau_1 - \mathbf{X}^2_{ij}\alpha}{\sigma_{\mathbf{W}_{2i}}}\big), \tag{3.6}$$

$$\begin{aligned} Pr[y_{ij} = d] &= Pr(\tau_d < y^*_{ij} < \tau_{d+1}) = Pr(\tau_d < \mathbf{X}^2_{ij}\alpha + \mathbf{W}_{2i} < \tau_{d+1}) \\ &= \Phi\big(\frac{\tau_{d+1} - \mathbf{X}^2_{ij}\alpha}{\sigma_{\mathbf{W}_{2i}}}\big) - \Phi\big(\frac{\tau_d - \mathbf{X}^2_{ij}\alpha}{\sigma_{\mathbf{W}_{2i}}}\big), \end{aligned} \tag{3.7}$$

$$Pr[y_{ij} = D] = Pr(y^*_{ij} > \tau_D) = Pr(\mathbf{X}^2_{ij}\alpha + \mathbf{W}_{2i} > \tau_D) = 1 - \Phi\big(\frac{\tau_d - \mathbf{X}^2_{ij}\alpha}{\sigma_{\mathbf{W}_{2i}}}\big) \tag{3.8}$$

where $\mathbf{W}_{2i} \sim N(0, \sigma^2_w)$, and Φ is the CDF of the standard normal distribution and $d = 1, , D - 1$. Let N be the total number of observations, and N_d be the number of individuals who select category d to indicate their watching indention. Then the likelihood for the intention model can be expressed as:

$$\begin{aligned} L(y_{ij}|\mathbf{U}_i) = &\Big[\Phi\big(\frac{\tau_1 - \mathbf{X}^2_{ij}\alpha}{\sigma_{W_{2i}}}\big)\Big]^{N_1} \left(\prod_{d=1}^{D-1} \Big[\Phi\big(\frac{\tau_{d+1} - \mathbf{X}^2_{ij}\alpha}{\sigma_{W_{2i}}}\big) - \Phi\big(\frac{\tau_d - \mathbf{X}^2_{ij}\alpha}{\sigma_{W_{2i}}}\big)\Big]^{N_d}\right) \\ &\Big[1 - \Phi\big(\frac{\tau_d - \mathbf{X}^2_{ij}\alpha}{\sigma_{W_{2i}}}\big)\Big]^{N_D}. \end{aligned} \tag{3.9}$$

3.3.2 Joint Model Specification

The joint model consists of the two sub-models described earlier, for the longitudinal emotion data and watching intention data. These two models are linked through stochastic dependence between $\mathbf{W}_{1i}(t)$ and $\mathbf{W}_{2i}$, which are the random effects that capture unobserved subject-specific characteristics. We include random effects for the intercept and the slope to account for the subject-specific emotion baseline and the speed of emotion change. Specifically, $\mathbf{W}_{1i}(t)$ and $\mathbf{W}_{2i}$ can be expressed as

$$\begin{aligned}\mathbf{W}_{1i}(t) &= u_{1i} + u_{2i}t \\ \mathbf{W}_{2i} &= v_1 u_{1i} + v_2 u_{2i} + u_{3i}.\end{aligned} \tag{3.10}$$

The random effects for the intercept and the slope are denoted u_{i1} and u_{i2}, and both are assumed to follow a Normal distribution: $\mathcal{N}(\mathbf{X}_i^3\gamma, \sigma_u)$, where $\mathbf{X}_i^3$ contains age and gender for each subject. The parameters v_1 and v_2 in the intention model measure the association between the two sub-models. The random term u_{i3} also follows a Normal distribution with mean 0. We assume that all measurements for a single individual are independent, conditional on the random effects $\mathbf{U}_i$. We can then write down the full likelihood of the joint model as:

$$p(\mathbf{H}_{ij}, y_{ij}, \mathbf{U}_i) = L(\mathbf{H}_{ij}|\mathbf{U}_i)L(y_{ij}|\mathbf{U}_i)p(\mathbf{U}_i) \tag{3.11}$$

where

$$L(\mathbf{H}_{ij}|\mathbf{U}_i) = \prod_{t=1}^{T} p(h_{ijt}|\mathbf{U}_i) \tag{3.12}$$

$$\begin{aligned}L(y_{ij}|\mathbf{U}_i) = &\Big[\Phi\big(\frac{\tau_1 - \mathbf{X}_{ij}^2\alpha}{\sigma_{W_{2i}}}\big)\Big]^{N_1} \left(\prod_{d=1}^{D-1}\Big[\Phi\big(\frac{\tau_{d+1} - \mathbf{X}_{ij}^2\alpha}{\sigma_{W_{2i}}}\big) - \Phi\big(\frac{\tau_d - \mathbf{X}_{ij}^2\alpha}{\sigma_{W_{2i}}}\big)\Big]^{N_d}\right) \\ &\cdot \Big[1 - \Phi\big(\frac{\tau_d - \mathbf{X}_{ij}^2\alpha}{\sigma_{W_{2i}}}\big)\Big]^{N_D}.\end{aligned} \tag{3.13}$$

We apply uninformative prior distributions on all parameters in the joint model. For all components in α, β, γ, ζ, τ, v_1 and v_2, a prior normal distribution $N(0, 100)$ is imposed. The inverse of a gamma distribution $Gamma(0.01, 0.01)$ is used for all variances σ.

3.3.3 Bayesian Variable Selection

To identify a parsimonious intention model, we need to select a limited set of predictors. However, using standard model tests based on the log-marginal density (LMD), the number of models to be considered is quite large. In particular, because we have $k = 45$, then the marginal likelihood of a total of $2^{45} = 3.52 \times 10^{13}$ models has to be estimated. Therefore, we need algorithms that efficiently search the model space. Researchers have proposed several MCMC methods for variable selection, including Kuo & Mallick's method ([139]), Gibbs Variable Selection (GVS) ([140]), Stochastic Search Variable Selection (SSVS) ([141]), Reversible jump Metropolis ([142]) and Adaptive shrinkage ([143, 144]).

In the watching intention model there are 45 predictors in total, including 24 aggregated emotion measurements. We apply Gibbs Variable Selection (GVS) developed by Dellaportas et al. [2000] [140] to estimate the posterior probability for all possible models ([145]). The goal is to find coefficients that are small enough to be insignificant and shrink them towards 0. In the GVS approach, the coefficients of the regression model are assumed to have prior distributions with a mixture of a point mass at 0 and a diffuse distribution elsewhere.

An auxiliary indicator variable I_k is introduced, with $I_k = 0$ indicating absence of the covariate k in the model and $I_k = 1$ indicating presence. The regression coefficient δ_k is then defined as the product of this indicator variable and the effect size parameter α_k: $\delta_k = I_k \alpha_k$. Thus, we have:

$$\delta_k = \begin{cases} 0 & \text{if } I_k = 0 \text{ (spike)} \\ \\ \alpha_k & \text{if } I_k = 1 \text{ (slab)} \end{cases} \tag{3.14}$$

where $k = 1, 2, ..., K = 45$.

The intention model for watching the movie in equation (3) can then be modified as follows:

$$y^*_{ij} = \mathbf{X}^2_{ij} \delta + u_{3i}. \tag{3.15}$$

The Gibbs variable selection method allows for dependence between α_k and I_k. Therefore, the joint density is calculated as $P(I_k, \alpha_k) = P(\alpha_k|I_k)P(I_k)$. The effect size parameter α_k is assumed to have a mixture prior: $P(\alpha_k|I_k) = (1 - I_k)N(\tilde{\mu}, \tau^2) + I_k N(0, \sigma^2)$, where $(\tilde{\mu}, \tau^2)$ requires tuning such that the values of α_k proposed when $I_k = 0$ is appropriate. The parameter σ^2 is the prior variance of α_k, which is fixed. A Bernoulli distribution $Bern(0.5)$ is imposed on the indicator I_k.

Posterior distributions of the parameters are obtained by Markov Chain Monte Carlo (MCMC). The model fitting is implemented utilizing JAGS. The MCMC runs are monitored using the package "runjags" implemented in the R-language[3]. We run two parallel MCMC chains with over-dispersed initial values. We discard the first 10000 iterations as burn-in, and use another 10000 iterations to obtain the posterior distribution of the parameters. The Gelman-Rubin diagnostic is used to check convergence of the model ([146]). Posterior mean, standard deviation (SD),

[3] http://cran.r-project.org

95% credible intervals and Bayesian p-values are calculated.

3.4 Experiment Results

3.4.1 Results on Bayesian Variable Selection

We perform Bayesian variable selection on the standardized data for the watching intention model. Table 3.2 shows the posterior mean and standard deviations for all the selection indicator variables I_k. If the frequency of the variable selection indicator in the posterior distribution $p(I_k|y_{ij})$ exceeds a preset cutoff, we consider including the k^{th} predictor in the X matrix into the intention model. Otherwise, it is excluded from the model. Before running the variable selection, the full GVS model consists of a total of 45 predictors. Among these 45 predictors, 24 are from the aggregated emotion measurements, such as the average level of emotion, the slope of emotion progress over time, emotion peak measurements and emotion levels in the first and last scene. Two participant-related variables, income and whether trailer influences the decision to watch a movie, are obtained from the questionnaire and included as well. We also added 19 trailer-related data consisting of scene, volume data from each trailer and movie information extracted form IMDb , such as scene number, average scene length, music volume measures, movie release time, star and director, as described above in Section 3.

Mean values of the indicator variables, which represent probabilities of corresponding variables to be selected in the model, are estimated (Table **??**). Some variables are almost never included (**Mean** < 0.05), such as `DisgustPeakIndex`,

Table 3.2: Bayesian variable selection results. The posterior means of indicators are reported. Estimates that are greater than the cutoff 0.2 are in bold.

Variable	Parameter	Mean	Variable	Parameter	Mean
DisgustPeakNum	I_1	**0.45450**	noVocalMean	I_{24}	**0.27370**
HappinessPeakNum	I_2	0.01560	volumeMean	I_{25}	0.01450
SurprisePeakNum	I_3	0.03930	DisgustPeak	I_{26}	0.05330
DisgustPeakDuration	I_4	0.04160	HappinessPeak	I_{27}	**0.58940**
HappinessPeakDuration	I_5	0.05680	SurprisePeak	I_{28}	0.18140
SurprisePeakDuration	I_6	**0.31020**	DisgustPeakIndex	I_{29}	0.01890
sceneNum	I_7	0.05300	HappinessPeakIndex	I_{30}	0.02490
DisgustAvg	I_8	0.03500	SurprisePeakIndex	I_{31}	0.03220
HappinessAvg	I_9	0.07000	DisgustEnd	I_{32}	0.01220
SurpriseAvg	I_{10}	0.13250	HappinessEnd	I_{33}	**0.47220**
DisgustCoef	I_{11}	0.01330	SurpriseEnd	I_{34}	0.06280
HappinessCoef	I_{12}	**0.78530**	DisgustStart	I_{35}	**0.98520**
SurpriseCoef	I_{13}	0.01010	HappinessStart	I_{36}	0.08930
volCoef	I_{14}	**0.84130**	SurpriseStart	I_{37}	0.03240
Influencetrailer	I_{15}	0.06540	VolumePeak	I_{38}	0.01260
Income	I_{16}	0.00710	VolumePeakInd	I_{39}	0.04310
SceneLenAvg	I_{17}	**0.21710**	VolumeEnd	I_{40}	0.02790
SceneLongestInd	I_{18}	**0.98350**	VolumeStart	I_{41}	**0.44040**
Summer	I_{19}	0.13700	MusicPeak	I_{42}	**0.31210**
Holiday	I_{20}	0.16380	MusicPeakInd	I_{43}	0.09840
Director	I_{21}	0.00650	MusicEnd	I_{44}	0.02560
Star	I_{22}	0.05240	MusicStart	I_{45}	**0.30030**
SoVocalCoef	I_{23}	**0.30560**			

HappinessPeakIndex, SurprisePeakIndex, HappinessPeakNum, SurprisePeakNum, SurpriseCoef, DisgustCoef, SurpriseStart, DisgustPeakDuration, DisgustAvg, DisgustEnd, Income, Director, VolumeMean, VolumePeak, VolumePeakInd, VolumeEnd and MusicEnd. This result suggests that factors including where the emotion peak is located at, how many positive emotion spikes, how much surprise change over time, all measurements of disgust expect the number of peaks, personal income, director of the movie, the average, peak and end of total volume and the music volume in the last scene do not affect the intention of watching a movie at all. Others have high probability of being included (**Mean** > 0.5), such as HappinessCoef, VolCoef, SceneLongestInd, HappinessPeak, Disguststart, which indicates the significant roles these variables might play in determining the watching intention.

To include most of the important factors while maintaining a reasonable size of the intention model, we choose 0.2 as the cutoff for variable selection in Table 3.2, and based on that decide to include the following 14 predictors in the intention model: the highest happiness level within the scene (HappinessPeak), number of emotion peaks that exceed 75% of the maximum value of disgust (DisgustPeakNum), duration of surprise level exceed 75% of the maximum (SurprisePeakDuration), slope of happiness level (HappinessCof), disgust level in the first scene (HappinessStart), happiness level in the end scene (HappinessEnd), the average scene length (SceneLenAvg), location of the longest scene (SceneLongestInd), slope of total volume (VolCoef), slope of trailer music (MusicCoef), average of music volume (MusicMean), the highest music volume within the scene (MusicPeak), and the average total and music volume in the first scene (VolumeStart, MusicStart).

3.4.2 Joint Model Selection

Using the model with 14 predictors, we test several nested versions of the model. We use the Deviance Information Criterion (DIC) as the model selection criteria. The DIC is defined as $\mathbf{DIC} = p_D + \bar{D}$. The expectation $\bar{D} = E^{\theta}[D(\theta)]$ measures how well the model fits the data, and the effective number of parameters is calculated as $p_D = \bar{D} - D(\bar{\theta})$, where the deviance $D(\theta) = -2\log f(y|\theta) + 2\log h(y)$ and $f(y|\theta)$ is the likelihood function for the observed data y given the parameter θ. As the number of parameters in the model increases, p_D also increases. However, $\bar{D}$ will decrease instead, which indicates better fit. Overall, the model with smaller total DIC is preferred.

Table 3.3 shows the DIC scores for joint models with different random effect terms $\mathbf{W}_{1i}(t)$ and $\mathbf{W}_{2i}$. We run two parallel MCMC chains with 10,000 burn-in iteration period and 10,000 iterations in JAGS for all models. The total DIC scores for all the joint models are calculated. The models differ in the random effects that are included, and that link the emotion and the watching intention model components. We examine 10 models in total. Model I is a separated model with no random effects at all. We then introduce a random intercept in $\mathbf{W}_{1i}(t)$ to build a separated model (Model II) and a joint model with association through the random intercept (Model III). The next seven models in the tables introduce a random slope term in $\mathbf{W}_{1i}(t)$. Model IV is a separate model in which emotions and intentions are independent. Association between the two sub-models is introduced through random intercept (Model V), random slope (Model VI) and both (Model VII). The final three models (Model VIII, IX and X) repeat the same pattern as Model V, VI and VII, with an extra random effect component in $\mathbf{W}_{2i}$. The last model (Model X) is the joint model with full complexity, while Model IV is the reduced model with no association between two sub-models.

In the sequel, we let

$$\mathbf{W}_1(t) = (\mathbf{W}_{11}(t), \mathbf{W}_{12}(t), ..., \mathbf{W}_{1N}(t))^T,$$

$$\mathbf{W}_2 = (\mathbf{W}_{21}, \mathbf{W}_{22}, ..., \mathbf{W}_{2N})^T,$$

$$\mathbf{u}_m = (u_{m1}, u_{m2}, ..., u_{mN})^T$$

where $m = 1, 2, 3$ and N is the total number of individuals in the analysis.

We start with Model I with no random effects in either of the two sub-models, which has the largest DIC score (432817.622). When we add separate random effects into both sub-models, this dramatically decreases the DIC score to 263807.815, as shown for Model II in Table 3.3. In Model III, we link $\mathbf{W}_{1i}(t)$ and $\mathbf{W}_{2i}$ through a common random intercept, which results in a slight increase in total DIC scores (263892.612). When we introduce the random effect for the time-slope in Model IV, this further decreases the DIC score to 260045.538. Models I to IV are independent models, in the sense that there are no random effects in the level or slope of the emotions that link the emotion model with the intention model. Associating the two sub-models through a common random intercept (Model V) or a random slope (Model VI) alone does not bring down the DIC score (260122.899 and 260154.729, respectively). Linking two sub-models through both random intercept and slope (VII) at the same time results in a slightly smaller DIC score (260122.823), which is

Table 3.3: Joint Model Selection Results.

Model	$\mathbf{W}_1(t)$	$\mathbf{W}_2(t)$	DIC
I	0	0	432817.622
II	u_1	u_3	263807.815
III	u_1	$v_1 u_1$	263892.612
IV	$u_1 + u_2 t$	u_3	260045.538
V	$u_1 + u_2 t$	$v_1 u_1$	260122.899
VI	$u_1 + u_2 t$	$v_2 u_2$	260154.729
VII	$u_1 + u_2 t$	$v_1 u_1 + v_2 u_2$	260122.823
VIII	$u_1 + u_2 t$	$v_1 u_1 + u_3$	260037.397
IX	$u_1 + u_2 t$	$v_2 u_2 + u_3$	260044.718
X	$u_1 + u_2 t$	$v_1 u_1 + v_2 u_2 + u_3$	260037.103

still higher than Model IV. However, if we keep the random effect in the intention model, but also link $\mathbf{W}_{1i}(t)$ and $\mathbf{W}_{2i}$ through the random intercept (Model VIII) or the slope (Model IX), the DIC score decreases to 260037.397 and 260044.718, respectively, which is a better fit compared to Model IV. This suggests that the association between two sub-models in random intercept and slope is important. The DIC achieves its lowest value (2660037.103) when the two models are associated by both random intercept and slope (Model X), although it is only slightly smaller than the DIC of model VII. We select Model X as our final joint model for emotion and intention data. To summarize, this model has 14 predictor variables, and has a linking random intercept and slope linking the longitudinal emotion data and intention data. It indicates the intention to watch a movie is related to two individual-specific emotion patterns, their initial emotion level and rate of emotion changes.

3.4.3 Comparison of the Joint Model and the Reduced Model

After selecting the best joint model, we compare the estimates of the reduced model (Model IV) and the joint model (Model X) (see Table 3.4). The joint model has smaller DIC (260037.103) compared to the reduced model (260045.538), which indicates better fit to the emotion and intention data. The posterior estimates for the reduced model and joint model are similar, with some minor differences in the effects and their P-values. Specifically for the emotion model, coefficient estimates and significances do not seem to change. In the intention model, the peak number of

Table 3.4: Posterior mean of parameters and Bayesian p-values for both the joint model and the reduced model.

		Joint model		Reduced model	
Variable	Parameter	Mean (SD)	P-value	Mean (SD)	P-value
`SubGenreMatch`	$\beta_{1,1}$	**0.033(0.004)**	0	**0.033(0.004)**	0
`Seentrailer`	$\beta_{1,2}$	0.008(0.006)	0.1757	0.008(0.006)	0.1818
`Seenmovie`	$\beta_{1,3}$	**0.157(0.008)**	0	**0.157(0.008)**	0
`Scene`	ζ_1	**-0.002(0)**	0	**-0.002(0)**	0
`Volume`	ζ_2	-0.087(0.055)	0.102	-0.09(0.053)	0.0947
`Music`	ζ_3	-0.133(0.084)	0.1209	-0.13(0.083)	0.123
`DisgustPeakNum`	α_1	**0.036(0.014)**	0.0088	**0.034(0.014)**	0.0158
`SurprisePeakDuration`	α_2	0.004(0.002)	0.08	**0.004(0.002)**	0.0495
`HappinessCoef`	α_3	3.894(9.967)	0.704	4.208(9.868)	0.6688
`VolCoef`	α_4	-1.867(9.952)	0.8482	-1.6(9.923)	0.8622
`SceneLenAvg`	α_5	0.005(0.004)	0.27	0.004(0.004)	0.3539
`SceneLongestInd`	α_6	**0.025(0.01)**	0.0081	**0.027(0.01)**	0.0075
`NoVocalCoef`	α_7	0.316(9.883)	0.9719	0.496(9.987)	0.9575
`NoVocalMean`	α_8	-12.061(6.608)	0.0664	**-12.7(6.531)**	0.0498
`HappinessPeak`	α_9	**2.012(0.659)**	0.0011	1.25(0.667)	0.0585
`HappinessEnd`	α_{10}	**4.19(0.762)**	0	**3.493(0.823)**	0
`DisgustStart`	α_{11}	**-10.753(3.41)**	0.0026	**-12.857(3.697)**	0
`VolumeStart`	α_{12}	-2.287(2.814)	0.4313	-2.419(2.882)	0.3636
`MusicPeak`	α_{13}	**6.305(2.691)**	0.0176	**6.598(2.752)**	0.0177
`MusicStart`	α_{14}	**-10.418(5.05)**	0.0368	-10.348(5.082)	0.0511
`Intercept`	γ_1	0.001(0.01)	0.8853	0.001(0.01)	0.8622
`Age`	γ_2	0(0)	0.8768	0(0)	0.8756
`Gender`	γ_3	0(0.004)	0.9769	0(0.004)	0.9632
`RandomIntercept`	v_1	**-0.821(0.184)**	0	–	–
`Randomslope`	v_2	1.504(9.744)	0.8743	–	–

disgust and location of the longest scene have very similar posterior estimates with more than 95% of the posterior mass away from zero. The posterior estimates of the happiness level in the end scene has a higher mean in the joint model compared to the reduced model, both with almost 100% of it posterior mass away from zero. The posterior mean estimate is higher for disgust level in the first scene and lower for music peak value, compared to the reduced model (all with more than 95% posterior mass away from zero).

Moreover, we observe differences in the intention model on the significance between the reduced model and the joint model. For example, the peak duration of surprise is "significant" in the reduced model, while the Bayesian p-value is 0.08 for the joint model and it would thus not be statistically significant there according to standard criteria. Similarly, the Mean music volume is significant in the reduced model, but only marginally so in the joint model; the posterior mean estimate of its' negative effect is also lower in the joint model. The happiness peak, however, has no statistically significant effect in the reduced model, but is strongly significant and higher estimated mean value in the joint model. Similarly, while music volume at the start of the trailer is marginally significant in the reduced model, it is significant in the joint model. But the posterior mean estimate is lower in the joint model. This shows that linking the emotion and intention components of the model results in statistically and substantively different findings.

In the longitudinal emotion model, the preference of comedy subgenre (`SubGenreMatch`) has a significant positive effect on emotion. If the individuals preference matches with the trailer subgenre, the happiness level of the viewer is higher. Previous exposure to the movie (`SeenMovie`) also plays a significant role in determining emotion. If an individual has watched the movie before, which suggests that their liking of the movie is already higher compared to movies they did not decide to watch, the happiness level while watching the trailer is higher. The happiness level generally decreases as the trailer continues. Total volume and music volume do not seem to have significant effects on emotion, which suggests that simply making the trailer louder does not increase audiences emotional engagement.

In the watching intention model, the peak and end of happiness level (`HappinessPeak` and `HappinessEnd`, respectively) both have positive effects on watching intention. This suggests that instead of the average level, it is the peak and end of a positive emotional experience that contribute to the watching intention afterwards. Also there is a significant negative effect of the disgust levels in the first scene on the watching intention, which suggests that a person that experiences a negative feeling the beginning of the trailer is unlikely to watch the movie. The total number of disgust peaks has a positive effect on watching intention, however. It could indicate that a mix of different types of emotion responses may increase the watching intention of a movie. When the longest scene is placed later in the trailer, the watching intention increases. Watching intention is also higher when the music is louder at its peak or when the music volume in the first scene is lower. Higher music volume at its peak will enhance the audiences response to the trailer. However, a starting scene with loud music does not seem to be appealing and will hurt the watching intention afterwards.

Finally, the effect of the random intercept on watching intention is negative and significant, which confirms the association between the emotion model and intention model. It suggests that the individual-specific baseline level (intercept) of happiness negatively affects the watching intention.

3.5 Discussion

In this paper, we use non-intrusive face-tracking systems to capture the moment-to-moment emotional response to comedy trailers and jointly analyze the longitudinal emotion model subject to dependent final watching intention model. We use a large set of explanatory variables to explain emotions and watching intentions, and a Bayesian variable selection model to test for variable inclusion that improves model fit. Two random effects representing the individual-specific baseline emotion level and emotion progression are introduced to link together these two models.

The first interesting finding in our analysis is that several emotion peak, end and starting values are selected by Bayesian variable selection to enter the watching intention model, while the average emotions are not. This result supports the Peak-end theory proposed by Fredrickson and Kahneman [1993] [147], which described the phenomena that people judge their experience predominantly by its most intense point and its end, not by the average or total sum of every moment of the experience. In our study, the peak and end of the pleasurable feeling of happiness have a positive effect on watching intention of comedy movies. The peak volume level of the music also was found to increase the watching intention.

The second discovery is that the start point of the experience also plays a role in people's overall evaluation of the experience. If an unpleasant feeling, such as disgust, is higher at the beginning, the watching intention decreases. Similarly, a loud music volume at the beginning might catch the audiences attention, but will actually decrease their watching intention for the movie. Moviegoers have complained about the movie trailers being "too loud" ([107]) and our findings confirms the negative effect of loud music at the beginning of the trailer on the watching intention of the movie.

We also find that the longest scene that could contain more details and allows the audiences to fully understand the story in the trailer is better placed towards the end. Perhaps surprisingly, more peaks in disgust have a positive effect on the watching intention. This might be due to the fact that disgust, even though an unpleasant feeling could be induced by different types of humor and jokes, such as slapstick, sarcasm, prejudices and "bathroom" humor, and more if this type of humor is appreciated by the audience.

By understanding the effects of different aspects of the design of comedy trailers on the success of the movies, studios can create movie trailers more efficiently by putting the most emotionally intense scenes in a better place and adjusting the music volume of the trailer. To summarize, we suggest filmmakers to consider the following while making trailers for comedy movies:

- At the beginning: avoid starting trailers with scenes or jokes that might induce negative feeling or make audiences uncomfortable, since bad first impression hurts watching intention for the movie. Also, scenes with lower music volumes are preferably used in the beginning.

- In the middle: include scenes with different types of humors if possible, for instance, slapstick, sarcasm and prejudices. One scene that has high music volume but within audiences comfort level should be included in the middle of the trailer.

- At the end: choose one of the funniest scenes from the movie and use it for the end of the trailer. It would be even better if the last scene were long enough to intrigue audiences.

In addition, we believe that using the joint model is beneficial for marketing research institutes or companies such as nViso to analyze longitudinal emotion response data and the overall evaluation including watching intention, rating and purchase intention.

Chapter 4: Conclusions and Future Research

In this dissertation we examine how to effectively design two types of online contents, e-commerce website and movie trailer videos. We first analyzed three eye-tracking experiments on static e-commerce websites to examine the joint effects of facial expression and gaze direction of models on viewers attention. Then we analyze a face-tracking study in comedy movie trailer videos and investigated the association between emotional responses and viewers watching intention. The findings in these two studies provide online retailers and filmmakers insights to attract and direct viewer's attention and emotion engagement.

In Chapter 2, we concluded that a model gaze, when directed at the product, can be used to orient viewers attention and that positive effect from happiness expression when a model looks at the viewer carries over to the product or brand, for both American and Chinese cultures. Due to the culture difference in using the eyes and mouth as cues to recognize and interpret smiles, a model that looks at the viewer with a happy expression draws more attention to the brand for American participants, while a model that looks at the product with a happy expression draws more attention to the brand for Chinese participants. The racial match between a model and the viewer also exacerbated the attention effects of facial expression.

These findings suggest that online retailers need to consider the impact of facial cues in relation to culture by adapting websites to users with different cultural background.

As with any research, the study in this chapter has limitations that suggest directions for future research. First, we focused only on cosmetic and perfume products in this study. A wider range of products could be investigated in future research to help generalize the present findings. Second, we only used female models in our experiments (the product categories studied were targeted at women). We did find effects of the gender of participants, although this was not the primary focus of the present research. The gender of the model employed in a website and its interaction with the gender of the viewer may be relevant in its own right and could be studied further. Third, we focused on two facial expressions which are common on e-commerce websites: neutral and happy, where especially we employed non-Duchenne smiles in this study. From the small content analysis of a sample of Western and Asian websites for apparel and cosmetics, it appeared that around half of the websites uses models with non-Duchenne smiles. Thus, although the non-Duchenne smile is very common, future research, could study Duchenne smiles, and other less common facial expressions, such as surprise.

In Chapter 3, we jointly analyzed the longitudinal emotional responses and dependent final watching intention. We confirmed the “Peak-end” effect of positive emotions and musics on the watching intention. We also found that a negative start point of the experience decrease the intention. A mixed types of emotion could potentially boost the overall experience.

There are also limitations this study that shed light on the future study, we analyze the content of movie trailer videos in two aspects: one is the scene cut, which was derived based on the similarity in visual image of the frames in the trailers. The second one is sound volume, including total volume and music volume. The advantage of these content variables is that they can automatically be derived using available software, without human intervention. However, one could also do a content analysis on the script of movie trailer, which might require natural language processing (NLP) or human editors. This would enable one to investigate questions such as, how smooth the story is in the trailer and how well the narratives go along with the music. Further study should look into these factors of movie trailers to give filmmakers a more complete suggestion on how to improve trailer design.

Another possible extension of this study is to use a follow-up study on actual movie going behavior of the individuals that participated in the study. Nevertheless, even though movie trailer is the main source of influencing movie-going behavior and was shown to have a positive and significant effect on expected revenues ([148]), a good trailer does not always guarantee the success of the movie. Long-term box office is determined by many other factors such as production cost, star power, volume and valence of online reviews, release time, sequels, genre, age-rating, competition and distribution company ([149, 150]), while some research has suggested that spending money on advertisements/trailers is effective only for high quality movies ([27]). These variables would need to be included in such a follow-up study.

We investigate the impact of emotional response to comedy trailers, which aims at provoking laughter and joy from the audiences. Future research could extend the present research to other movie genres, including action, drama, documentary, horror, music, sci-fi and fantasy, etc. Emotions other than happiness should then also be focused on. For example, fear is the most effective emotional response during exposure to horror movies. Whether our conclusions on which variables affect the production of an effective trailer for a comedy movie applies to other movie genres as well, would require further investigation.

Finally, future research could also investigate other dimensions of responses to movie trailers that cannot be measured by facial muscle movement, such as arousal, which could be measured by skin conductance (SC) and heart rate (HR) and attention, which could be captured by eye-tracking. We hope the present study provides a good starting point for such future research.

Overall, in these two eye-tracking and face-tracking studies, we investigate viewer's attention and emotional responses to online stimuli. We learn the joint effect of model gaze direction and facial expression on attention in a static e-commercial website in the first study and the effect of moment-to-moment emotional responses on watching intentions of video trailers in the second study. We gave constructive suggestions to online retailers and filmmakers on how to make effective websites and comedy trailers, respectively. The methods applied in these two studies can be extended to investigate a much wider range of online contents in marketing research.

Bibliography

[1] M. Wedel and P. Pieters. *Looking at vision, eye/face/head tracking of consumers for improved marketing decisions, Chapter 9, In: Routledge Companion on Future of Marketing, L. Moutinho, E. Bign, and A.K. Manrai (eds.).* Rouledge, 2014.

[2] M. Wedel and P. Pieters. *Eye-Tracking for Visual Marketing. Foundations and Trends in Marketing.* Now Publishers Inc, 2008.

[3] P. Ekman, W. V. Friesen, and J. C. Hager. Facial action coding system: The manual. *A Human Face*, 2002.

[4] T. Teixeira, M. Wedel, and R. Pieters. Emotion induced engagement in internet video ads. *Journal of Marketing Research*, 49(2):144–159, 2012.

[5] T. Teixeira, R. Picard, and R. Kaliouby. Why, when, and how much to entertain consumers in advertisements? a web-based facial tracking field study. *Marketing Science*, pages 1–19, 2014.

[6] nViso. Say it to my face! applying facial imaging to understanding consumer emotional response. *AMSRS Conference 2011*, 2011.

[7] Y. W. Seo, S. W. Chae, and K. C. Lee. The impact of human brand image appeal on visual attention and purchase intentions at an e-commerce website. *in Intelligent Information and Database Systems: Springer Berlin Heidelberg*, pages 1–9, 2012.

[8] K. Hassanein and M. Head. Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping. *International Journal of Human – Computer Studies*, 65(8):689–708, 2007.

[9] D. Cyr, M. Head, H. Larios, and B. Pan. Exploring human images in website design: a multi-method approach. *MIS Quarterly*, 33(3):539–566, 2009.

[10] K. Humphrey and G. Underwood. The potency of people in pictures: Evidence from sequences of eye fixations. *Journal of Vision*, 10(10):1–10, 2010.

[11] S. R. H. Langton, A. S. Law, A. M. Burton, and S. R. Schweinberger. Attention capture by faces. *Cognition*, 107(1):330–342, 2008.

[12] L. Xiao and M. Ding. Just the faces: Exploring the effects of facial features in print advertising. *Marketing Science*, 33(3):338–352, 2014.

[13] A.D. Engell and J.V. Haxby. Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia*, 45:3234–3241, 2007.

[14] A. Frischen, A. P. Bayliss, and S. P. Tipper. Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, 4(133):694–724, 2007.

[15] M. F. Mason, E. P. Tatkow, and C. N. Macrae. The look of love : Gaze shifts and person perception. *Psychological Science*, 16(3):236–239, 2005.

[16] C. K. Friesen and A. Kingstone. The eyes have it! reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 5(3):409–495, 1998.

[17] S.B. Hutton and S. Nolte. The effect of gaze cues on attention to print advertisements. *Applied Cognitive Psychology*, 25(887-892), 2011.

[18] P. Ekman and E. L. Rosenberg. What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (facs, 2nd ed.). *New York, NY, US: Oxford University Press*, 2005.

[19] H. Cho and S. Norbert. When good pictures make for good products: Consumer misattribution effects in virtual product presentation environments. *Advances in Consumer Research*, 33(1):637–638, 2006.

[20] D. A. Small and N. M. Verrochi. The face of need: Facial emotion expression on charity advertisements. *Journal of Marketing Research*, 46(6):777–787, 2009.

[21] R. B. Jr. Adams and R. E. Kleck. Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science*, 14:644–647, 2003.

[22] J. K. Hietanen and J. M. Leppanen. Does facial expression affect attention orienting by gaze direction cues? *Journal of Experimental Psychology*, 29(6):227–228, 2003.

[23] C. Blais, R. E. Jack, C. Scheepers, D. Fiset, and R. Caldara. Culture shapes how we look at faces. *PLoS ONE*, 3(8):e3022, 2008.

[24] R. E. Jack, R. Caldara, and P. G. Schyns. Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology*, 141(1):19–25, 2012.

[25] W. J. Qualls and D. J. Moore. Stereotyping effects on consumers evaluation of advertising: Impact of racial differences between actors and viewers. *Psychology & Marketing*, 7(2):135–151, 1990.

[26] R.J. Faber and T.C. OGuinn. Effect of media advertising and other sources on movie selection. *Journalism Quarterly*, 61(2):371–377, 1984.

[27] J. Vanderhart and S.N. Wiggins. Advertising, quality, and motion picture performance. *Mimeo, Texas A&M University*, 2004.

[28] J. Last. Opening soon. *Wall Street Journal*, May 2004.

[29] PwC. Pwc's 2013 entertainment report. *The Hollywood Reporter*, 2013.

[30] H. Baumgartner, M. Sujan, and D. Padgett. Patterns of affective reactions to advertisements: The integration of moment-to-moment responses into overall judgments. *Journal of Marketing Research*, Vol. XXXIV:219–232, May 1997.

[31] J. Elpers, M. Wedel, and R. Pieters. The influence of moment-to-moment pleasantness and informativeness on zapping tv commercials: A functional data and survival analysis approach. *Advances in Consumer Research*, 29, 2002.

[32] S.K. Hui, T. Meyvis, and H. Assael. Analyzing moment-to-moment data using a Bayesian functional linear model: Application to TV show pilot testing. *Marketing Science*, pages 222–240, January 2014.

[33] B. He and S. Luo. Joint modeling of multivariate longitudinal measurements and survival data with applications to parkinson's disease. *Statistical Methods in Medical Research*, NA:1–13, April 2013.

[34] A. Tsiatis and M. Davidian. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica*, 14:809–834, 2004.

[35] P. Wang, W. Shen, and M.E. Boye. Joint modeling of longitudinal outcomes and survival using latent growth modeling approach in a mesothelioma trial. *Health Serv Outcomes Res Methodol.*, 12:182–199, June 2012.

[36] M. Bindemann, A. M. Burton, and S. R. H. Langton. How do eye gaze and facial expression interact? *Visual Cognition*, 16(6):708–733, 2008.

[37] A. Senju and T. Hasegawa. Direct gaze captures visuospatial attention. *Visual Cognition*, 12(1):127–144, 2005.

[38] M. Bindemann, A. M. Burton, I. T C. Hooge, R. Jenkins, and E. F. D. Haan. Faces retain attention. *Psychonomic Bulletin & Review*, 12(6):1048–1053, 2005.

[39] J. Driver, G. Davis, P. Ricciardelli, P. Kidd, E. Maxwell, and S. Baron-Cohen. Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, 6(5):509–540, 1999.

[40] C. K. Friesen, C. Moore, and A. Kingstone. Does gaze direction really trigger a reflexive shift of spatial attention? *Brain & Cognition*, 57(1):66–69, 2005.

[41] A. Kingstone, C. Tipper, J. Ristic, and E. Ngan. The eyes have it! an fmri investigation. *Brain and Cognition*, 55(2):269–271, 2004.

[42] S. R. Langton, R. J. Watt, and I. Bruce I. Do the eyes have it? cues to the direction of social attention. *Trends in Cognitive Sciences*, 4(2):50–59, 2000.

[43] U. Dimberg and M. Thunmerg. Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, 39:39–45, 1998.

[44] E. Hatfield, J. T. Cacioppo, and R. L. Rapson. Emotional contagion. *New York: Cambridge University Press*, 1994.

[45] S. Blairy, P. Herrera, and U. Hess. Mimicry and the judgment of emotional facial expressions. *Journal of Nonverbal Behavior*, 23(1):5–41, 1999.

[46] U. Dimberg, M. Thunberg, and K Elmehed. Unconscious facial reactions to emotional facial expressions. *Psychological Science*, 11(1):86–89, 2000.

[47] M. Stel and A. V. Knippenberg. The role of facial mimicry in the recognition of affect. *Psychological Science*, 19(10):984–985, 2008.

[48] E. Fox, V. Lester, R. Russo, R. J. Bowles, Pichler A., and K. Dutton. Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition & Emotion*, 14(1):61–92, 2000.

[49] E. Fox, R. Russo, and K. Dutton. Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition & Emotion,*, 16(3):355–379, 2002.

[50] R. G. Coss, S. Marks, and U. Ramakrishnan. Early environment shapes the development of gaze aversion by wild bonnet macaques (macaca radiata). *Primates*, 43(3):217–222, 2002.

[51] K. A. Nichols and B. G. Champness. Eye gaze and the gsr. *Journal of Experimental Social Psychology*, 7(6):623–626, 1971.

[52] L. Ewing, G Rhodes, and E. Pellicano. Have you got the look? gaze direction affects judgements of facial attractiveness. *Visual Cognition*, 18(3):321–330, 2010.

[53] L. Schilbach, A. M. Wohlschlaeger, N. C. Kraemer, A. Newen, N. J. Shah, G. R. Fink, and K. Vogeley. Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, 44(5):718–730, 2006.

[54] D. Sander, D. Grandjean, S. Kaiser, T. Wehrle, and K. R. Scherer. Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European Journal of Cognitive Psychology*, 19(3):470–480, 2007.

[55] C. A. Conway, Jones B. C., L. M. DeBruine, and A. C. Little. Evidence for adaptive design in human gaze preference. *Proceedings of the Royal Society B: Biological Sciences*, 275(1630):63–69, 2008.

[56] B. C. Jones, L. M. DeBruine, A. C. Little, C. A. Conway, and D. R. Feinberg. "integrating gaze direction and expression in preferences for attractive faces. *Psychological Science*, 17(7):588–591, 2006.

[57] A. P.and Frischen A. Bayliss, M. J. Fenske, and S. P. Tipper. Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, 47(1):14–27, 2007.

[58] H. Ozono, M. Watabe, and S. Yoshikawa. Effects of facial expression and gaze direction on approachavoidance behaviour. *Cognition & Emotion*, 26(5):943–949, 2012.

[59] P. Ekman and W. V. Friesen. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2):124–129, 1971.

[60] B. Mesquita and N. H. Frijda. Cultural variations in emotions: A review. *Psychological Bulletin*, 112(2):179–204, 1992.

[61] M. Yuki, W. W. Maddux, and T. Masuda. Are the windows to the soul the same in the east and west? cultural differences in using the eyes and mouth as cues to recognize emotions in japan and the united states. *Journal of Experimental Social Psychology*, 43(2):303–311, 2007.

[62] R. E. Jack, C. Blais, C. Scheepers, P.G. Schyns, and R. Caldara. Cultural confusions show that facial expressions are not universal. *Current Biology*, 29:15431548, 2009.

[63] P. Ekman, R. J. Davidson, and W. V. Friesen. The duchenne smile: emotional expression and brain physiology ii. *Journal of Personality and Social Psychology*, 58(2):342–353, 1990.

[64] X. Mai, Y. Ge, L. Tao, H. Tang, C. Liu, and Y. J. Luo. Eyes are windows to the chinese soul: Evidence from the detection of real and fake smiles. *PLoS ONE*, 6(5):e19903, 2011.

[65] L. L. Kontsevich and C. W. Tyler. What makes mona lisa smile? *Vision Research*, 44(13):1493–1498, 2004.

[66] J. M. Leppänen and J. K. Hietanen. Is there more in a happy face than just a big smile? *Visual Cognition*, 15(4):468–490, 2007.

[67] P. G. Schyns, L. Bonnar, and F. Gosselin. Show me the features! understanding recognition from the use of visual information. *Psychological Science*, 13(5):402–409, 2002.

[68] M. L. Smith, G. W. Cottrell, F. Gosselin, and P. G. Schyns. Transmitting and decoding facial expressions. *Psychological Science*, 16(3):184–189, 2005.

[69] H. Ozono, M. Watabe, S. Yoshikawa, S. Nakashima, N. O. Rule, N. Ambady, and Jr. R. B. Adams. What's in a smile? cultural differences in the effects of smiling on judgments of trustworthiness. *Letters On Evolutionary Behavioral Science*, 1(1):15–18, 2010.

[70] M. H. Bond and H Komai. Targets of gazing and eye contact during interviews: Effects on japanese nonverbal behaviour. *Journal of Personality and Social Psychology*, 34(1276-1284), 1976.

[71] H. D. Pierson and M. H. Bond. How do chinese bilinguals respond to variations of interviewer language and ethnicity? *Journal of Language and Social Psychology*, 1:123–139, 1982.

[72] M. LaFance and C. Mayo. Cultural aspects of nonverbal communication. *International Journal of Intercultural Relations/Communication*, 2(71-89), 1978.

[73] C. R. Huston, E. D'Ouville, and T. H. Willis. Race and gender effects on consumer identification of products with celebrity endorsers. *Marketing Management Journal*, 13(1):85–99, 2003.

[74] D. Byrne, G. L. Jr. Clore, and P. Worchel. Effect of economic similarity-dissimilarity on interpersonal attraction. *Journal of Personality and Social Psychology*, 4(2):220–224, 1966.

[75] H. Tajfel and J.C. Turner. The social identity theory of intergroup behavior. *in Psychology of Intergroup Relations, S. Worchel and W.G. Austin, Chicago, IL, USA: Nelson-Hall*, pages 7–24, 1986.

[76] D. M. Messick and D. M. Mackie. Intergroup relations. *Annual Review of Psychology*, 40(1):45–81, 1989.

[77] J. W. Cagley and R. N. Cardozo. White response to integrated advertising. *Journal of Advertising Research*, 10(2):35–39, 1970.

[78] M. J. Schlinger and J. T. Plummer. Advertising in black and white. *Journal of Marketing Research*, 9(4):149–153, 1972.

[79] R. Deshpande and D. M. Stayman. A tale of two cities: Distinctiveness theory and advertising effectiveness. *Journal of Marketing Research*, 31(1):57–64, 1994.

[80] T. E. Whittler. Viewers processing of actors race and message claims in advertising stimuli. *Psychology & Marketing*, 6(4):263–286, 1989.

[81] T. E. Whittler and J. DiMeo. Viewers' reactions to racial cues in advertising stimuli. *Journal of Advertising Research*, 31(6):37–46, 1991.

[82] O. Appiah. The effectiveness of "typical-user" testimonial advertisements on black and white browsers' evaluations of products on commercial websites: Do they really work? *Journal of Advertising Research*, 47(1):14–27, 2007.

[83] T. Elias, O. Appiah, and G. Li. Effects of strength of ethinic identity and product presenter race on black consumer attitudes: a multiple-group model approach. *Journal of Interactive Advertising*, 11(2):13–29, 2011.

[84] I. Kareklas and M. Polonsky. A meta-analytic review of racial similarity effects in advertising. *Advances in Consumer Research*, 37:829–832, 2010.

[85] K. Krmer, G. Bente, S. Luo, U. J. Pfeiffer, S. Han, and K. Vogeley. Influence of ethnic group-membership and gaze direction on the perception of emotions. a cross-cultural study between germany and china. *Plos One*, 8(6):e6635, 2013.

[86] L. M. Brown, M. M. Bradley, and P.J. Lang. Affective reactions to pictures of ingroup and outgroup members. *Biological Psychology*, 71(3):303–311, 2006.

[87] M. Wedel and R. Pieters. Eye fixations on advertisements and memory for brands: A model and findings. *Marketing Science*, 19(4):297–312, 2000.

[88] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. V. D. Weijer. Eye tracking: A comprehensive guide to methods and measures. *Oxford, UK: Oxford University Press.*, 2011.

[89] P. J. A. Unema, S. Pannasch, M. Joos, and B. M. Velichkovsky. Time course of information processing during scene perception: The relationship between scccade amplitude and fixation duration. *Visual Cognition*, 12(3):473–494, 2005.

[90] S. Pannasch and B. M. Velichkovsky. Distractor effect and saccade amplitudes: Further evidence on different modes of processing in free exploration of visual images. *Visual Cognition*, 17(6/7):1109–1131, 2009.

[91] J. Simola, J. Kuisma, A. Öörni, L. Uusitalo, and J. Hyönä. The impact of salient advertisements on reading and attention on web pages. *Journal of Experimental Psychology: Applied*, 17(2):174–190, 2011.

[92] R. Pieters, M. Wedel, and R. Batra. The stopping power of advertising: Measures and effects of visual complexity. *Journal of Marketing*, 74(5):48–60, 2010.

[93] M. M. Bradley, L. Miccoli, M. A. Escrig, and P. J. Lang. The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45(4):602–607, 2008.

[94] A. L. Chasteen, D. C. Burdzy, and J. Pratt. Thinking of god moves attention. *Neuropsychologia*, 48:627–630, 2010.

[95] D. G. Gozli, A. L. Chasteen, and J. Pratt. The cost and benefit of implicit spatial cues for visual attention. *Journal of Experimental Psychology*, 142(4):1028–1046, 2013.

[96] S. C. Goodhew, B. McGaw, and E. Kidd. Why is the sunny side always up? explaining the spatial mapping of concepts by language use. *Psychonomic Bulletin & Review*, 21(5):1287–1293, 2014.

[97] M. B. Holbrook. Beyond attitude structure: Toward the informational determinants of attitude. *Journal of Marketing Research*, 15(4):545–556, 1978.

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

[99] C. P. Puto and W. D. Wells. Informational and transformational advertising: the differential effects of time. *Advances in Consumer Research*, 11(1):638–643, 1984.

[100] M. M. Bradley, M Codispoti, B. N. Cuthbert, and P. J. Lang. Emotion and motivation i: Defensive and appetitive reactions in picture processing. *Emotion*, 1(3):276–298, 2001.

[101] R. Goldstein, L. C. Walrath, J. A. Stern, and B. D. Strock. Blink activity in a discrimination task as a function of stimulus modality and schedule of presentation. *Psychophysiology*, 22(6):629–635, 1985.

[102] J. A. Stern, L. C. Walrath, and R. Goldstein. The endogenous eyeblink. *Psychophysology*, 21(1):22–23, 1984.

[103] J. R. Hauser, G. L. Urban, G. Liberali, and M. Braun. Website morphing. *Marketing Science*, 28(2):202–223, 2009a.

[104] J. R. Hauser, G. L. Urban, G. Liberali, and M. Braun. Response to comments on "website morphing.". *Marketing Science*, 28(2):227–228, 2009b.

[105] S.T. Eastman, D.E. Bradbury, and R.S. Nemes. Influences of previews on movie goers expectations. *Current research in film: Audiences, economics, and law*, 1:51–57, 1985.

[106] J. Wasko. *How Hollywood works.* London: Sage, 2003.

[107] L. Kerman. *Coming attractions: Reading American movie trailers.* Texas film and media studies series. Austin: University of Texas Press, 1 edition, 2004.

[108] C.B. Stapleton. Mixed reality and experiential movie trailers: Combining emotions and immersion to innovate entertainment marketing. In *Proceedings of 2005 international conference on human-computer interface advances in modeling and simulation (SIMCHI05*, pages 23–27, 2005.

[109] S.S. Jones. Trailer music: A look at the overlooked. *The Oxford Student*, 2011.

[110] B.R. Litman. Predicting success of theatrical movies: An empirical study. *Journal of Popular Culture*, 16(9):159–175, 1983.

[111] B.R. Litman and A. Kohl. Predicting financial success of motion pictures: The 80's experience. *The Journal of Media Economics*, 2(1):35–50, 1989.

[112] R. Sharda and D. Delen. Predicting box-office success of motion pictures with neural networks. *Expert Systems with Applications*, 30(2):243–254, 2006.

[113] J. Eliashberg and S. M. Shugan. Film critics: Influencers or predictors? *Journal of Marketing*, 61(68-78), 1997.

[114] S. A. Ravid. Information, blockbusters, and stars: A study of the film industry. *Journal of Business*, 72(4):463–492, 1999.

[115] B. A. Austin. Portrait of an art film audience. *Journal of Communication*, 34:74–87, 1984.

[116] B. A Austin and T. F Gordon. Movie genres: Toward a conceptualized model and standardized definition. *In Current Research in Film: Audiences, Economics and the Law*, 3, 1987.

[117] C. Levene. Marketing art-films to college students. *Working Paper, The Wharton School, University of Pennsylvania.*, 1992.

[118] J. Eliashberg, S. K. Hui, and Z. J. Zhang. From story line to box office: A new approach for green-lighting movie scripts. *Management Science*, 53(6):881–893, 2007.

[119] J. Hewig, D. Hagemann, J. Seifert, M. Gollwitzer, E. Naumann, and D. Bartussek. Brief report: A revised film set for the induction of basic emotions. *Cognition and Emotion*, 19(7):1095–1109, 2005.

[120] C.R. Plantinga and G.M. Smith. *Passionate views: Film, cognition, and emotion.* Johns Hopkins University Press, 1999.

[121] A.M. Isen and R. Patrick. The effect of positive feelings on risk taking: When the chips are down. *Organizational Behavior and Human Performance*, 31(2):194–202, 1983.

[122] J.S. Lerner, D.A. Small, and G. Loewenstein. Heart strings and purse strings: Carry-over effects of emotions on economic decisions. *Psychological science*, 15(5):337–341, 2004.

[123] J.S. Lerner and D. Keltner. Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion*, 14(4):473–493, 2000.

[124] M. Lewis, J.M. Haviland-Jones, and L.F. Barrett. *Handbook of emotions.* Guilford Press, 2010.

[125] J.K. Burgoon and S.B. Jones. Toward a theory of personal space expectations and their violations. *Human Communication Research*, 2:131–146, 1976.

[126] A. P. McGraw and Warren C. Benign violations: Making immoral behavior funny. *Psychological Science*, pages 1–9, Aug. 2010.

[127] I.B. Mauss and M.D. Robinson. Measures of emotions: A review. *Cognition and Emotion*, 23(2):209–237, February 2009.

[128] M.D. Robinson and G.L. Clore. Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality & Social Psychology*, 1(83):198–215, 2002.

[129] A. Öhman, A.and Hamm and K. Hugdahl. *Cognition and the autonomic nervous system: Orienting, anticipation, and conditioning. In: Cacioppo, JT.; Tassinary, LG.; Berntson, GG., editors. Handbook of psychophysiology.* New York: Cambridge University Press, 2000.

[130] A.M. Dale and M.I. Sereno. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience*, 5(2):162176., 1993.

[131] J.A. Detre and T.F. Floyd. Functional MRI and its applications to the clinical neurosciences. *Neuroscientist*, 11(1):86–89, 2000.

[132] N.D. Volkow, B. Rosen, and L. Farde. Imaging the living human brain: Magnetic resonance imaging and positron emission tomography. *Proceedings of National Academy of Sciences of the United States of America*, 94(7):2787–2788, 1997.

[133] J.A Bachorowski and M.J. Owren. Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychological Science*, 6(4):219224, 1995.

[134] R. Adolphs. Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, 1(1):21–62, 2002.

[135] P. Ekman and W. Friesen. *Facial action coding system: A technique for the measurement of facial movement.* Consulting Psychologists Press, Palo Alto, 1978.

[136] P.J. Bradley, M.M.; Lang. *Measuring emotion: Behavior, feeling, and physiology. In: Lane, RD.; Nadel, L., editors.* New York: Oxford University Press, 2000.

[137] R. Gertner. *Motion Picture Almanac.* New York: Quigley Publishing Co., 1981.

[138] S. Fischoff. Race and sex differences in patterns of film attendance and avoidance. *APA Convention, Los Angeles, August 15.*, 1994.

[139] L. Kuo and B. Mallick. Variable selection for regression models. *SankhyāGeMc93 : The Indian Journal of Statistics*, 60:65–81, 1998.

[140] P. Dellaportas, J.J. Forster, and I. Ntzoufras. Bayesian variable selection using the Gibbs sampler, 2000.

[141] E.I. George and R.E. McCulloch. Variable selection via Gibbs sampling. *Journal of American Statistical Association*, 88:881–889, 1993.

[142] P. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82:711–732, 1995.

[143] S. Xu. Estimating polygenic effects using markers of the entire genome. *Genetics*, 163:789–801, 2003.

[144] M. A. T. Figueiredo. Adaptive sparseness for supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:1150–1159, 2003.

[145] R.B. O'Hara and M.J. Sillanpää. A review of Bayesian variable selection methods: What, how and which. *Bayesian Analysis*, 4(1):85–118, 2009.

[146] A. Gelman and D.B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–511, 1992.

[147] B. L. Fredrickson and D. Kahneman. Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, 65(1):45–55, 1993.

[148] A. Elberse and B. Anand. The effectiveness of pre-release advertising for motion pictures: An empirical investigation using a simulated market. *Information Economics and Policy*, 19:319–343, 2007.

[149] P. Topf. Examining success at the domestic box-office in the motion picture industry. *Honors Projects Paper 110, Economics Department at Digital Commons, Illinois Wesleyan University*, 2010.

[150] C. Dellarocas, X. Zhang, and N.F. Awad. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, 21(4):23–45, 2007.