Emotions Don’t Lie: Introducing an Artificial Intelligence Tool to Measure Ad Effectiveness

Patrick Kremer
Goethe University Frankfurt
Benjamin Abdel-Karim
Goethe University Frankfurt
Patrick Felka
Goethe University Frankfurt
Ju-Young Kim
Goethe-Universität Frankfurt
Oliver Hinz
Goethe University Frankfurt

Cite as:
Kremer Patrick, Abdel-Karim Benjamin, Felka Patrick, Kim Ju-Young, Hinz Oliver (2021), Emotions Don’t Lie: Introducing an Artificial Intelligence Tool to Measure Ad Effectiveness. *Proceedings of the European Marketing Academy*, 50th, (93722)

Paper from the 50th Annual EMAC Conference, Madrid, May 25-28, 2021

EMOTIONS DON'T LIE: INTRODUCING AN ARTIFICIAL INTELLIGENCE TOOL TO MEASURE AD EFFECTIVENESS

Abstract

The role of emotions in advertising has long been studied. Yet, in-depth insights into how emotions develop throughout the ad exposure and the effects on ad effectiveness are scarce. To address this research gap we develop a new AI-based tool to measure consumer emotions in real time utilizing readily available smartphone sensors for inexpensive scalability of this solution. The findings of our study suggest that emotional ads increase recall over a rather monotonous and unemotional ad. Furthermore, the results hint on the fact that variations of emotions throughout the ad exposure elevate recall over emotional ads that mainly focus on one specific emotional target (e.g. happiness). Concluding, we highlight the importance of this novel tool for theory as well as for practitioners which become enabled to rely on data gathered in the field via smartphone sensors instead of artificially constructed copy tests.

Keywords: Emotions, Advertising, Artificial Intelligence

Track: Methods, Modelling & Marketing Analytics

1. Introduction

"What does science really know about advertising? What kind of ads are most likely to make us buy products? The true and unsatisfying answer is that there is a lot we just don't know", the Atlantic stated in 2011 regarding the role of emotions in advertising (Thompson, 2011). Prior research suggests that emotions play a crucial role in the effectiveness of an ad. Positive emotional tones tend to increase the attitude towards an ad as well as its corresponding effectiveness (Eckler & Bolls, 2011; Tucker, 2015). For instance, positive and emotion-focused ads have been shown to drive ad sharing on multiple social media platforms (Tellis, MacInnis, Tirunillai, & Zhang, 2019). Dobele, Lindgreen, Beverland, Vanhamme, and van Wijk (2007) find that emotions foster the success of viral marketing campaigns. Specifically, they find surprise combined with further emotional reactions to increase the ad effectiveness. These findings emphasize the role of emotions in advertising; however, insights about how the interaction of different evoked emotions may affect consumers or insights about how to manage effective timings of emotional ads are scarce.

To measure ad effectiveness, companies usually apply specific copy testing procedures such as the day-after recall in order to evaluate if consumers remember ad contents (Arnold & Bird, 1982; Mehta & Purvis, 2006). Companies use these testing procedures both, before broadcasting the ads to reduce risks of ad failure as well as while ads are "live" in the field to track their success. The disadvantages of these procedures are, however, that they are very costly in money and time (Hamelin, Moujahid, & Thaichon, 2017) and that consumers are surveyed in a very artificial manner as they are usually recruited in a group of participants watching the ad in unfamiliar surroundings (Stewart, Pechmann, Ratneshwar, Stroud, & Bryant, 1985).

In this study, we introduce a new, AI-based tool to forecast recall and measure ad effectiveness that can be applied by marketers in the field through utilizing state-of-the-art, built-in smartphone sensors to detect consumer emotions multiple times per second. Using this novel approach, traditional copy testing becomes obsolete, saving substantial resources for marketers. Furthermore, using smartphone sensors for emotion detection enables marketers to track emotions and shifts thereof in real-time throughout the exposure to the ad. This in turn allows for a deeper investigation of emotional reactions, variances in emotions throughout the exposure and the interplay of various emotions to better understand ad effectiveness. As such, our new test is able to apply emotion detection in a scalable, inexpensive manner in the natural, real-world environment of a consumer.

In the remainder of this paper, we provide insights into the past research on ad effectiveness and the role of emotions in advertising. We then present our new emotion detection tool based on readily available smartphone sensors. Afterwards, we investigate how consumer emotions vary throughout ad exposures and what the consequences are for ad recall. We conclude with a discussion of our theoretical as well as managerial contributions and avenues for future research in this realm.

2. Theoretical Background

Emotional states and their impact on consumer decision making have long been studied. Prior research unfolds the multitude of ways in which emotions influence the way we behave. As such, it has been shown that emotions play a crucial role in advertising. For instance, Williams and Drolet (2005) find that with increasing age consumers recall emotional ads to a greater extent. Geuens, Pelsmacker, and Faseur (2011) find that the effectiveness of emotional advertising is dependent on the product category which is consistent with a majority of prior research (Holmes & Crocker, 1987). Teixeira, Wedel, and Pieters (2012) examined the impact of surprise and joy on a viewer's perception of an ad. They find that surprise increases the attention strongly while joy leads to a higher retention of the viewer, hinting at the importance of intertwining various, distinct emotions. Similarly, Dobele et al. (2007) investigate consumers' sharing behavior of viral marketing campaigns and find that surprise is a crucial element of successful campaigns. Linking the surprising element to specific emotions that match the mood and intent of the shown ad (e.g. joy, sadness, anger) leads consumers to respond to the sent message more frequently, resulting in a more active sharing behavior.

The above examples show the importance of investigating various emotions throughout an ad exposure in order to capture the processes triggered by individual emotions as well as their interaction. Research largely agrees on the fact that the line between these different emotions is blurred and oftentimes mixed emotions occur throughout a single experience (Ruth, Brunel, & Otnes, 2002), such as feeling happy and sad in a touching moment. Furthermore, by watching a television ad the consumer will mostly experience a shift of emotions throughout the exposure, for instance from boredom to excitement (Rossiter & Percy, 2017). Nevertheless, most research either measures a single emotion at a given moment or continuously throughout the exposure. To address this challenge, we introduce a novel approach of emotion detection based on readily available sensors within modern mobile devices. The subsequently described emotion recognition tool is utilized to investigate real-time data of consumers exposed to various advertisements in order to create a deeper understanding of the effect of different stimuli on consumer emotions.

3. Methodology

For emotion recognition on a mobile device, we developed a mobile application for Apple iOS devices, which aims to capture and analyze facial information in real-time. For this purpose, the app has two main functions, the collection of training data that build the basis for our emotion classifier and the classification of collected data, which will be used during the lab experiment (next chapter) to classify emotions. In the following, we will describe the technical functionality of our app and the training data collection procedure in more detail.

Our iOS application is based on the open-source project Loki[1], which makes use of Apple's ARKit framework. This framework offers a variety of tools that analyze the scene visible to the device camera using computer vision techniques, such as face tracking or object detection. An essential tool we used from the ARKit framework for face recognition is the so-called AR-FaceAnchor. This tool allows the ARKit to track the facial movements in real-time, while simultaneously providing coefficients that represent detected facial expressions, the so-called blend shape coefficients. These coefficients are updated at a high frequency, approximately 10-20 times per second, which allows our application to track the facial feature movements in real-time. A well-known application that uses these coefficients is Apple's Animojis, which are basically animated emojis that are tracked to facial expressions of the user. Since these coefficients already describe facial expressions, they form the ideal base for the actual emotion recognition. In more detail, these coefficients are basically a vector of 51 individual coefficients representing the facial expressions, where each coefficient is a floating-point number that indicates the current position of that feature relative to its neutral configuration, ranging from 0.0 (neutral) to 1.0 (maximum movement). In general, these coefficients can be categorized into four primary areas: The left and right eye, the mouth and jaw, the eyebrows, and the cheeks and nose. Our application uses these

[1] https://github.com/nwhacks-loki/loki.

coefficients and accesses the front camera to detect the user's facial expression for data collection or emotion classification, while the user is looking at the screen.

To collect the needed training data, we developed a frontend for our mobile application, where users can record their facial expressions and label their corresponding emotions directly on the phone. In practice, the user starts the application, activates the camera in the application, and is asked which emotion he is currently showing. After the emotion is selected using a selection menu by the user, the data vector with the emotion label is sent to the server to train the model offline, to meet the demand respectively resources for the model training phase. Here the recorded and transited data samples are used as training data, where each vector represents the 51 coefficients or input features used for the Deep Learning (DL) model. More concrete, we developed and trained a feed-forward network. For the data collection procedure itself, we used multiple mobile devices to create a large dataset with many different variations of the emotions to train the model. Each time one of the devices transmitted new training data, the server restarted the training procedure and sent back an updated model to the devices. After the training phase, the final model was subsequently used for emotion recognition and in course of this study.

4. Study

We use the above-described model to investigate how various emotions impact the recall of different ads. Emotional ads have been shown to be capable to increase ad effectiveness, specifically when different emotions emerge throughout the ad exposure (Teixeira et al., 2012).

4.1 Method

To investigate this underlying mechanism in more detail, we recruited a sample of 63 student participants (41.7% female, median within the age group 18-22) in a laboratory experiment (data collection is still running; thus, the results are preliminary). Each participant was shown four advertisements with varying foci (the ads aimed to stimulate happiness, sadness, surprise and boredom) resulting in a set of 235 individual ad exposures and a set of 124,004 data points of tracked emotions.

The participants were seated in front of an iPhone X. After answering initial questions regarding demographics the participants watched the four different ads in a randomized order. The ads were selected through a pretest to generate as much difference in emotional reactions as possible. In the pretest, 30 participants (63.3% female, median age 24) saw a set of ten videos and were asked what emotion the video primarily generated for them (happy, sad, surprised, neutral). The videos with the clearest response across participants were selected for the subsequent study. As such, the selected videos showed a road safety video with a motorcycle accident to target surprise and three different advertisements of different brands. These were Edeka, a German supermarket chain, targeting sadness and Haribo, a German fruit gum company, targeting happiness. Further, as a fourth video the participants saw the ad of Thermacare, a brand producing heat patches. This last ad is rather neutral without much variation in emotions. The aim of incorporating this ad as well was to identify if unemotional videos generate recall in a different manner than emotional videos do. All participants were called a day after the experiment and asked what ads and contents they remember (unaided recall).

4.2 Results and Discussion

The results of this study are preliminary and will be accompanied by more in-depth analyses as soon as data collection is completed. We first investigate the results of the study graphically. Figure 1 shows the shares of emotions that have been detected across all participants over the time of the individual ad. Visualizing these differences already shows the power of our emotion detection tool such, that it is capable of showing emotional changes over time.

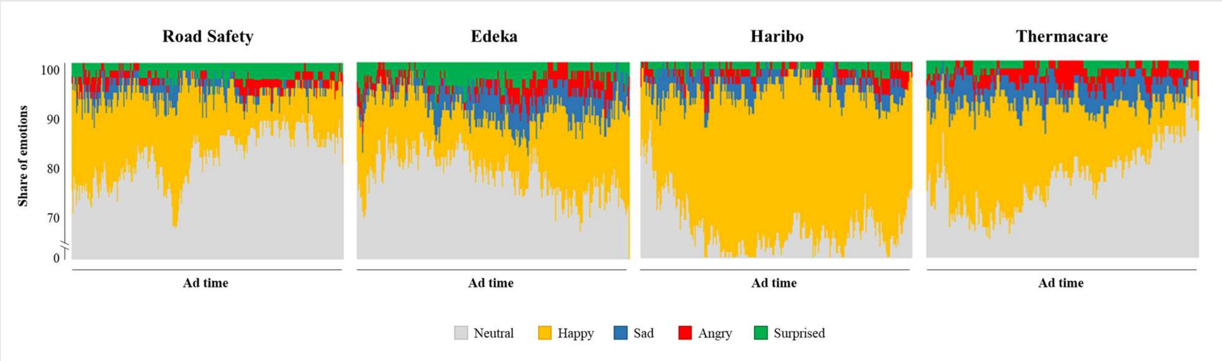

Figure 1: Share of emotions across participants per video over ad time

The visualization shows that different ads trigger different emotions over time. Interestingly, in the first moments of each ad the shares of detected emotions are fairly similar. Yet, over the course of the ad, depending on the ad content these values change. Table 1 shows the different characteristics of each ad.

Ad name	**Focal emotion in pretest (self-response)**	**Ad characteristics**
Road Safety	Surprise	• Twisting plot • Monotonous, with intense, short shock moment: Motorcycle accident
Edeka	Sadness	• Twisting plot • Starting from sad story turning into a touching end
Haribo	Happiness	• Monotonous plot • Light mood throughout the whole ad without plot turns
Thermacare	Boredom / Neutral	• Monotonous plot • Increasingly turns into scientific explanations

Table 1: Ad characteristics

The road safety video is rather monotonous except for a very short moment in which a car surprisingly hits a motorcycle after a third of the ad duration. Figure 1 shows a steep increase of emotional reactions of the participants within this time window. In contrast, the Edeka ad relies on a continuously twisting plot that is rather comprised as a story. An old man is faking his death in order to get his family to visit him for Christmas. The turn in the plot happens in the last third of the ad. Reviewing participants' reactions it is clearly visible that emotions are evoked by the plot. The individual participants get sad, angry or happy – all components of a touching moment. Compared to these two rather twisting plots, the Haribo ad shows high shares of emotional reactions, specifically happiness, yet, not that much variation throughout the video is visible. Finally, the Thermacare ad starts with a light mood but at the half of the spot turns into very scientific arguments with a doctor explaining graphs that show the effect of the heat patches. At this point the neutral emotion substantially gains shares indicating that the ad is getting more boring over time for the participants.

Yet, not only can we see how participants are reacting to the different ads but we can also refer these patterns to the recall rates for the different ads. As such, figure 2 shows the mean recall for the different advertisements. While only 40 percent of participants recalled the rather boring Thermacare ad, 75 percent of participants remembered the touching and twisting Edeka ad ($M_{Recall\_Edeka} = .75$, $M_{Recall\_Thermacare} = .40$; $p < .01$). Also, the recall of the road safety spot was significantly higher compared to the Thermacare ad ($M_{Recall\_Road\ Safety} = .61$, $M_{Recall\_Thermacare} = .40$; $p = .022$).

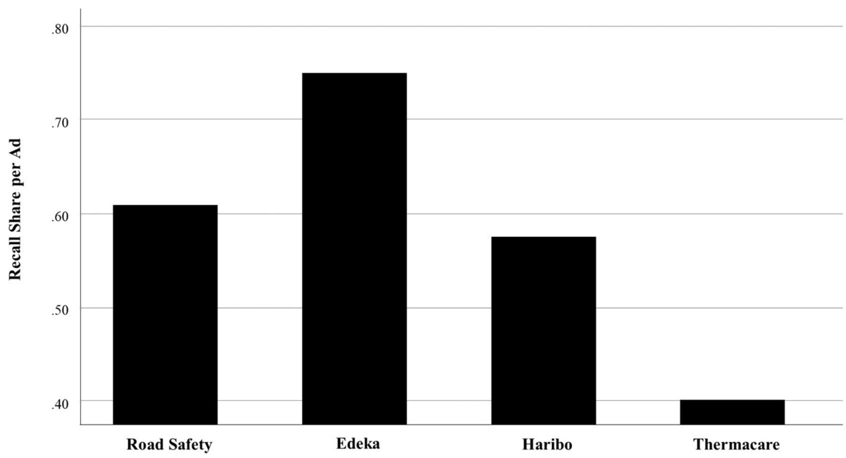

Figure 2: Recall rates per ad

Incorporating these findings, it seems that the more emotions are fluctuating over the course of an ad the better the participants can remember the content and recall the ad in general. While in the Edeka ad the participants really go through a journey of sadness and joy, in the road safety spot participants get shocked with a very intense but short moment of surprise. Both plots help the participants to recall the respective videos while an unemotional ad like the Thermacare spot does not foster recall as much. Ultimately, the Haribo ad is different in a sense that it evokes happy emotions for many participants. Yet, although more emotions increase the share of recall compared to the Thermacare ad, it is not as effective as an ad with changing emotional phases.

5. Conclusion and Limitations

The present study introduces a new tool that enables marketers to generate deep insights into the effectiveness of an ad on individuals. Based on artificial intelligence and readily available smartphone sensors this tool can be rolled out at scale without massive investments. It further allows to not only track recall rates of different advertisements but to elaborate the mechanisms behind this effect in depth. Aspects such as emotion shifts or mixed emotions can be investigated in order to tailor ads to the individual spectator. As such, the preliminary results of this study largely contribute to the research on the role of emotions in advertising while enabling marketers to transfer theoretical findings into practice.

In a subsequent step, the results of this study will be further investigated regarding the interactions of different emotions over time. Besides, the impact of the variation of emotions on recall will be analyzed in order to identify which mechanisms contribute to the effectiveness of an ad. Additionally, not only recall overall but also brand recall will be investigated to identify if there are details that are triggered more or less through emotional ads. Ultimately, future studies need to consider different product categories as well in order to identify if emotional ads have the same impact regardless of the nature of the product (e.g. utilitarian vs. hedonic).

Overall, the Atlantic seems to be still right: there is a lot we do not know about the role of emotions in advertising, however, with this new tool we may come one step closer.

6. References

Arnold, S. J., & Bird, J. R. (1982). The Day-after Recall Test of Advertising Effectiveness: A Discussion of the Issues. *Current Issues and Research in Advertising*, *5*(1), 59–68.

Dobele, A., Lindgreen, A., Beverland, M., Vanhamme, J., & van Wijk, R. (2007). Why pass on viral messages? Because they connect emotionally. *Business Horizons*, *50*(4), 291–304.

Eckler, P., & Bolls, P. (2011). Spreading the Virus. *Journal of Interactive Advertising*, *11*(2), 1–11.

Geuens, M., Pelsmacker, P. de, & Faseur, T. (2011). Emotional advertising: Revisiting the role of product category. *Journal of Business Research*, *64*(4), 418–426.

Hamelin, N., Moujahid, O. E., & Thaichon, P. (2017). Emotion and advertising effectiveness: A novel facial expression analysis approach. *Journal of Retailing and Consumer Services*, *36*, 103–111.

Holmes, J. H., & Crocker, K. E. (1987). Predispositions and the comparative effectiveness of rational, emotional and discrepant appeals for both high involvement and low involvement products. *Journal of the Academy of Marketing Science*, *15*(1), 27–35.

Mehta, A., & Purvis, S. C. (2006). Reconsidering Recall and Emotion in Advertising. *Journal of Advertising Research*, *46*(1), 49–56.

Rossiter, J. R., & Percy, L. (2017). Methodological Guidelines for Advertising Research. *Journal of Advertising*, *46*(1), 71–82.

Ruth, J. A., Brunel, F. F., & Otnes, C. C. (2002). Linking Thoughts to Feelings: Investigating Cognitive Appraisals and Consumption Emotions in a Mixed-Emotions Context. *Journal of the Academy of Marketing Science*, *30*(1), 44–58.

Stewart, D. W., Pechmann, C., Ratneshwar, S., Stroud, J., & Bryant, B. (1985). Methodological and Theoretical Foundations of Advertising Copytesting: A Review. *Current Issues and Research in Advertising*, *8*(2), 1–74.

Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-Induced Engagement in Internet Video Advertisements. *Journal of Marketing Research*, *49*(2), 144–159.

Tellis, G. J., MacInnis, D. J., Tirunillai, S., & Zhang, Y. (2019). What Drives Virality (Sharing) of Online Digital Content? The Critical Role of Information, Emotion, and Brand Prominence. *Journal of Marketing*, *83*(4), 1–20.

Thompson, D. (2011, October 27). Thinking vs. Feeling: The Psychology of Advertising. Retrieved from https://www.theatlantic.com/business/archive/2011/10/thinking-vs-feeling-the-psychology-of-advertising/247466/

Tucker, C. E. (2015). The Reach and Persuasiveness of Viral Video Ads. *Marketing Science*, *34*(2), 281–296.

Williams, P., & Drolet, A. (2005). Age-Related Differences in Responses to Emotional Advertisements. *Journal of Consumer Research*, *32*(3), 343–354.