Few and Far Between: Identifying Measures of Advertising Visuals that Correlate with Neural Engagement and Sales

Samuel B. Barnett, Northwestern University, USA
Moran Cerf, Northwestern University, USA

ABSTRACT

We present a novel approach to assess visual attention in advertisements. We find that movie trailers with fewer and larger attention-grabbing visuals generate higher collective neural engagement (i.e., Cross-Brain Correlation) among consumers in a movie theater. The number and size of attentional targets also correlate with future population-level ticket sales.

INTRODUCTION

With decreasing attention spans and increasing communication urgency, every second of an engaged mind counts even more. Content creators in every profession aim to engage their audiences in limited time, and the question of how to deliver a message quickly is especially pronounced in advertising. Advertisers and neuroscientists are constantly striving for better answers to this question, especially given the strong economic incentives to communicate efficiently. For context, a typical 30-second advertising slot during a primetime television news broadcast cost $5,000 in 2016; when more viewers are watching, such as during contentious political debates or sporting events, airtime rates often exceed that price per second (Poggi 2015). During *Super Bowl LI*, a single second of advertising time cost over $167,000 according to the American Marketing Association (Watters 2017). Furthermore, these costs only reflect the airtime and do not account for the cinematography and production expenses. Producing just one second of a Hollywood film or advertisement often costs over $20,000 ($22,109 mean budget per second for the ten highest grossing films in 2016; IMDb.com, Seattle, Washington).

Since images often convey messages more efficiently than words (Simion & Shimojo 2006; Cerf, Frady, and Koch 2009), visual attention to natural stimuli has increasingly become a focus of research. Certain commonalities in visual attention have been identified through a variety of techniques, such as eye tracking (Teixeira et al., 2010), saccade models (Itti & Koch 2001; Mackay et al. 2008), electroencephalography (EEG; Charland et al. 2013), functional magnetic resonance imaging (fMRI; Falk, Berkman, and Lieberman 2012; Poldrack 2008), and biometric measures (Ohme et al. 2011). In particular, elevated neural similarity across viewers during rich visual stimuli (e.g., videos) has been linked to heightened collective brain engagement, attention, memory, and population-level sales (Hasson et al. 2004, 2008; Barnett & Cerf 2015, 2017).

Recent research has shown that simpler advertisements with fewer words and lower visual entropy are linked to elevated neural similarity and engagement with content (Putrevu, Tan, and Lord 2004; Pieters, Wedel, and Batra 2010; Barnett et al. 2016).

By extension, we hypothesize that advertisements utilizing images with fewer attentional targets will promote greater neural similarity. However, it is not clear how the size of an attentional target will impact neural similarity. On one hand, a small visual might pinpoint collective attention (i.e., cause everyone to focus on a precise location within the image), thus increasing stimulus processing consistency. On the other hand, perceptual load theory suggests that an image with a larger attention-grabbing portion (e.g., high load) will minimize task-irrelevant attentional allocation (Lavie & Tsal 1994; Lavie 1995). Under this model, attentional targets are parsed as distinct entities (e.g., characters, faces, objects) rather than locations, so neural similarity is presumably increased if the audience simply focuses on the same entity (even if each viewer's gaze falls on a different location within that entity).

Put simply, does an image make brains think alike if it makes everyone look at the same location (as with a smaller attentional target, which minimizes gaze variance) or pay attention to the same entity (as with a larger attentional target, which minimizes distraction by other entities in the image)? We hypothesize that the latter is true; in other words, we expect that the average size of individual attentional targets throughout an advertisement will correlate with its neural similarity.

We address these hypotheses with a combination of neural and subjective report data. In a field experiment, we measure neural similarity as Cross-Brain Correlation (CBC) of moviegoers in a theater during natural viewing of trailers for upcoming feature films. Since movie trailers are dynamic advertising stimuli that both entertain and persuade, they have been the focus of many recent studies (cf., Boksem & Smidts 2015). We also perform an online survey to assess the number and location of attentional targets for images extracted from each second of the movie trailers.

Finally, in addition to showing that fewer and larger attentional targets correlate with higher CBC, we provide evidence that these metrics predict future sales.

METHODS

A total of 611 subjects collectively participated in this research; 59 viewed movie trailers in a theater while undergoing EEG recordings, and 552 analyzed images extracted from each second of these movie trailers.

EEG Field Experiment

Subjects and Procedure

Subjects watched trailers and movies while undergoing EEG recordings at a commercial theater (AMC Theatres, Northbrook, Illinois). Subjects selected a movie that they had not previously seen and were offered free admission in exchange for participation. To minimize non-neural artifacts in these recordings, subjects refrained from all other activities (e.g., drinking, eating, talking, moving) except viewing the screen. Subjects were fitted with an electrode cap and conductive gel was placed at each electrode site. All subjects were native English speakers with normal (or corrected-to-normal) vision and hearing who provided informed consent.

Task

After the setup procedure, subjects viewed the movie trailers and feature film naturally. Following the movie, subjects responded to a written questionnaire that asked them which trailers they recalled and other questions about their preferences.

EEG Data Acquisition

Subjects' neural data were collected using 32-channel EEG systems (Brain Products, Gilching, Germany) at a rate of 250 Hz. The electrode sites were distributed across the entire scalp according to the actiCAP 64Ch Standard-2 (green holders) montage.

Computation of Neural Similarity

We computed CBC, moment-to-moment synchrony in EEG data across subjects experiencing the same audiovisual stimuli, through-

out each movie trailer. At each electrode site, we measured neural activity over time as the power (dB) of alpha oscillations (Berger 1929), which are often associated with visual attention (Klimesch 2012, Dmochowski et al. 2014). We performed a Short-Time Fourier Transform (STFT) of the EEG signal at each timestep, filtered the Power Spectral Density (PSD) matrix, and multiplied the common logarithm (base 10) of the PSD matrix by 10. Next, we assembled a time series of activity at the given electrode site that was correlated with the stimulus-matched time series of activity at the corresponding site for each subject. At each timestep and site, we averaged the correlations for every pair of subjects. Lastly, we averaged across the sites to arrive at a single value of neural similarity at each timestep, thus producing the CBC time series. CBC values were normalized to range from zero (minimum) to one (maximum).

Stimuli. Each movie was typically preceded by six or seven movie trailers. Throughout the study, a total of 13 trailers (1,775 seconds in length) were presented more than once and subsequently recalled by more than one subject. On average, the length of each trailer was 136 ± 20 seconds (*mean $\pm$ standard deviation*).

Identification of Attentional Targets

Subjects and Procedure

We recruited subjects on a crowdsourcing Internet marketplace (Amazon Mechanical Turk, Seattle, Washington) to analyze still images extracted from the first frame of each second of the 13 movie trailers (i.e., 1,775 images), and 25 unique subjects responded to each image.

Task

Each image was divided into a grid of 60 locations (six rows, ten columns) and subjects were asked to (1) identify the location that "grabbed their attention the most," which we denote as their "primary attentional target," and (2) report the "total number of items that drew their attention" (e.g., characters, objects, graphics). Subjects spent 13 ± 8 seconds per image. We also included control tasks to verify reasonable effort; in these randomly interspersed tasks, subjects were instructed to choose an asterisked location in the image.

Data was discarded for subjects who failed to answer any control task accurately.

Measures of Central Tendency

We calculated the mean and median numbers of attentional targets across the 25 responses per image. Across all images in a given trailer, we calculated the mean of these means (i.e., "average mean") and mean of these medians (i.e., "average median").

### RESULTS

Movie trailer CBC averaged $.49 \pm .03$. The highest and lowest trailer CBCs corresponded with z-scores of 2.19 (*X-Men: Days of Future Past*) and -1.37 (*Mr. Peabody and Sherman*), respectively. The CBC distribution is right tailed; that is, engaging, high CBC trailers are rarer than would be expected in a symmetric normal distribution (positive skew of .59).

Most often, one attentional target was reported per image (37% of responses); the median was two (26%), and the mean was 2.12 ± 1.64. Zero attentional targets were reported for images with no discernible visual content (5%). Responses with more than ten reported targets (2%) were discarded as outliers for all of the following analysis. Also, the plurality (as well as median and mean) of primary attentional targets were in one of the four centermost locations of the frame (rows 3/4 × columns 5/6); however, this was only observed in 30% of responses.

Neural similarity tended to be higher for movie trailers with fewer attentional targets. We observed strong negative correlations between CBC and both average median ($r = -.67$, $p = .01$) and average mean ($r = -.64$, $p = .02$) numbers of attentional targets per image. The highest CBC trailer had an average median and average mean of 1.42 and 1.84 targets, respectively, which is significantly fewer ($p < .01$, unpaired two-sample t-tests) than the corresponding average median (2.00) and average mean (2.29) of the lowest CBC trailer.

Additionally, we observed that CBC was correlated with spatial dispersion of primary attentional target locations (see legend in Figure 1). In other words, for a given attention-grabbing visual, CBC increased as primary attentional target locations were further apart
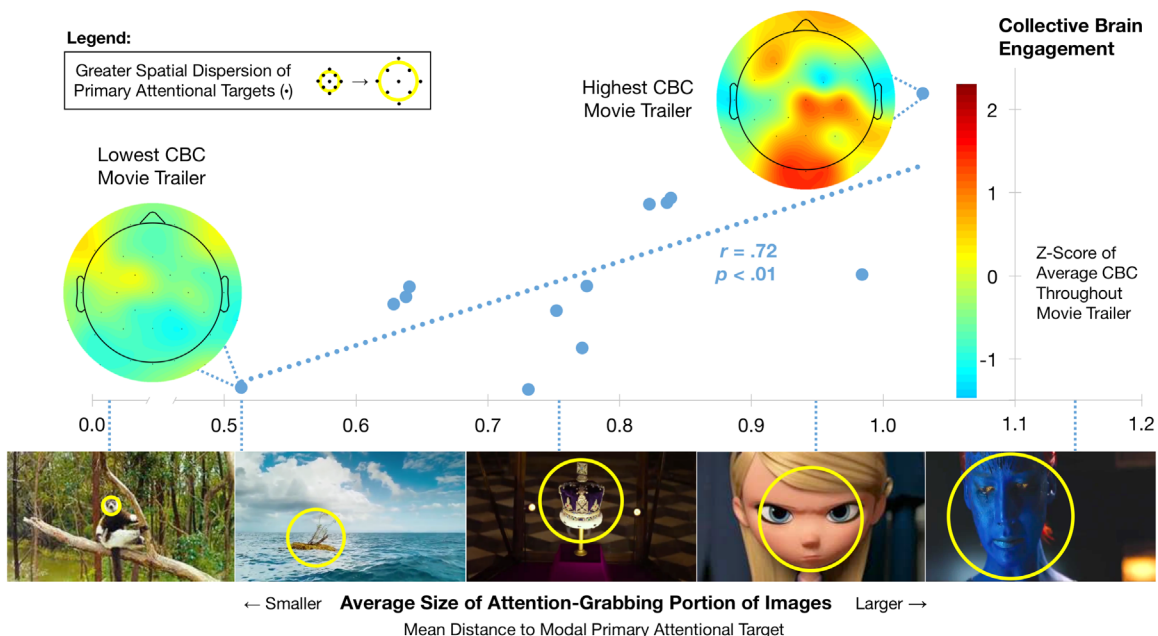
Figure 1: Size of Attention-Grabbing Visual Predicts Collective Brain Engagement During Movie Trailers.

from each other. Specifically, CBC was highly correlated with the standard deviations of primary attentional target locations along both the horizontal ($r = .80$, $p < .01$) and vertical ($r = .72$, $p < .01$) dimensions. Therefore, elevated CBC arose for attentional targets that were "few and far between": fewer items that drew subjects' attention, but primary targets "far between" each other.

To investigate this phenomenon further, we analyzed images with a median of one attentional target. For each image, we computed the mean distance of all subjects' primary attentional targets to the within-image mode of these responses, approximating the size of the attention-grabbing portion of the image. These averages of mean distance across trailers ranged from .51 to 1.03 units (where the distance between horizontally or vertically adjacent locations was one unit, and the distance between diagonally adjacent locations was Ö2 = 1.41 units). We found a strong positive correlation between CBC and average distance ($r = .72$, $p < .01$). That is, given a single attentional target, CBC increased with the size of the attention-grabbing area of the image (see Figure 1).

The standard deviation above/below average CBC throughout a given movie trailer (y-axis) is plotted against the average size of the attention-grabbing portion of images in that trailer (x-axis). For reference, five images are each shown with a dotted line indicating its attention-grabbing portion's size. Additionally, superimposed yellow circles delineate the attention-grabbing regions; the legend illustrates how the circle's size depends on the spatial dispersion of the attentional targets. Overhead EEG spatial maps (Niedermeyer & Lopes Da Silva 1993) of the regional CBC levels are shown for the lowest and highest CBC trailers (ranging from blue to red for lower and higher CBC, respectively; see gradient adjacent to vertical axis). In addition to greater average CBC, the highest CBC trailer has especially high CBC in posterior regions (represented in red) associated with visual processing areas of the brain.

Similarly, we measured the percentage of primary attentional targets that were at or adjacent to the within-image modal location (i.e., proportion of responses ≤ 1.41 units away from mode selection). These "near-mode percentages" were relatively high across trailers, ranging from 77.91% to 90.11%. Lower percentages corresponded with larger attentional areas since fewer primary attentional targets are concentrated near the modal location. Again, we found that CBC is directly related to the size of an attention-grabbing visual; CBC was negatively correlated with near-mode percentages ($r = -.71$, $p < .01$).

Given that CBC is positively correlated with the size of the attention-grabbing entity and negatively correlated with the number of attentional targets, we may gain additional predictive power by combining these metrics. For example, the fraction of (mean distance to within-image mode) ÷ (average median number of attentional targets) is more strongly correlated with CBC ($r = .79$, $p < .01$) than either the numerator or denominator alone.

In addition to being strongly correlated (positively or negatively) with CBC, these visual components are predictive of the eventual sales of the advertised movies (see Summary Table for statistics on each measure and its correlations with CBC and weekly sales).

The mean and standard deviation of each visual measure utilized in this work were calculated across trailers. The per-trailer averages for each measure were also compared to the average CBCs and ultimate weekly ticket sales of the full movie corresponding with each trailer, and the table displays the resulting correlations. All $|r| > .40$ were statistically significant ($p < .10$).

DISCUSSION

Fewer and larger attentional targets throughout movie trailers were significantly predictive of both higher CBC during viewing and increased sales of the advertised films. These findings agree with and extend prior literature on visual attention in advertising.

For example, lower numbers of attention-grabbing visuals have been associated with improved effectiveness of print advertisements (Book and Schick 1997; Aitchison 2012). Since consumers control the amount of time spent viewing a print advertisement, potential attentional targets are necessarily competing with each other for time and interest, so advertisers provide fewer degrees of freedom to increase focus on key items. We show this phenomenon also holds with video advertisements, despite that consumers do not control the presentation rate. In other words, competition for visual attention is continuously occurring (even frame by frame during a movie trailer).

Furthermore, when an item dominates more of the visual field, it reduces the set of other hypothetical targets for stray attentional resources. Also, when an item is larger, more visual details (e.g., colors, edges, textures) are visible, which demands greater within-item feature processing (i.e., higher perceptual load), thereby enhancing visual attention (Pieters, Wedel, and Batra 2010). Therefore, our observation that larger visuals generate higher brain engagement may be due to maximization of task-relevant (and minimization of task-irrelevant) attention according to perceptual load theory (Lavie & Tsal 1994; Lavie 1995).

Our efforts to capture natural attention in a commercial context were accompanied by many limitations. We were restricted to the theater's available films and their corresponding trailers; our findings may not hold for other types of video advertisements. Additionally, to maximize the authenticity of the movie viewing experience, subjects did not wear eye tracking goggles, and we instead supplemented the neural data with subjective reports of attentional targets in the online survey. These limitations can be addressed in future studies with different exploratory priorities.

Nonetheless, if "a picture is worth a thousand words" (Flanders 1911), then perhaps one with a single, large attentional target is worth a million.

Summary Table: Fewer and Farther-Distributed Attentional Targets Linked to Higher Collective Brain Engagement and Sales.

Visual Component	Measure	Mean ± Standard Deviation	Correlation with	
			CBC	Weekly Sales
Number of	Average Median	1.79 ± .22 targets	-.67	-.48
Attentional Targets	Average Mean	2.15 ± .19 targets	-.64	-.51
Primary Attentional	Near-Mode Proportion	84.75% ± 3.97%	-.71	-.50
Target Location	Standard Deviation (Horizontal)	15.94% ± 2.48% of width (10 units)	.80	.46
(Size of Attention-	Standard Deviation (Vertical)	14.37% ± 1.74% of height (6 units)	.72	.31
Grabbing Portion)	Mean Distance to Mode	.77 ± .14 units	.72	.39
Combined	(Mean Distance)/(Average Median)	.44 ± .13 units/target	.79	.53

REFERENCES

Aitchison, J. (2012), *Cutting Edge Advertising: How to Create the World's Best Print for Brands in the 21st Century*, FT Press.

Barnett, S. B. & Cerf, M. (2015), "Connecting on Movie Night? Neural Measures of Engagement Differ by Gender," *Advances in Consumer Research*, *43*, 314-318.

_____ (2017), "A Ticket for Your Thoughts: Method for Predicting Content Recall and Sales Using Neural Similarity of Moviegoers," *Journal of Consumer Research*, *44*(1), 160-181.

Barnett, S. B., White, H. M., & Cerf, M. (2016), "Keep It Simple Stimuli: Brain-Vetted Elements of Movie Trailers Predict Opening Weekend Ticket Sales," *Advances in Consumer Research*, 44, 285-289.

Berger, H. (1929). "Über das Elektrenkephalogramm des Menschen," *European Archives of Psychiatry and Clinical Neuroscience, 87*(1), 527-570.

Boksem, M. A., & Smidts, A. (2015), "Brain Responses to Movie Trailers Predict Individual Preferences for Movies and Their Population-Wide Commercial Success," *Journal of Marketing Research, 52*(4), 482-492.

Book, A. C., & Schick, D. C. (1997), *Fundamentals of copy & layout,* McGraw Hill Professional.

Cerf, M., Frady, E. P., & Koch, C. (2009), "Faces and Text Attract Gaze Independent of the Task: Experimental Data and Computer Model," *Journal of Vision, 9*(12).

Charland, P., Leger, P.-M., Allaire-Duquette, G., & Gingras, G. (2013), "A Neuroscientific Perspective For Assessing Student Engagement in E-Learning," *Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2013*, 1795-1800.

Dmochowski, J. P., Bezdek, M. A., Abelson, B. P., Johnson, J. S., Schumacher, E. H., & Parra, L. C. (2014), "Audience Preferences are Predicted by Temporal Reliability of Neural Processing," *Nature Communications, 5*.

Falk, E. B., Berkman, E. T., & Lieberman, M. D. (2012), "From Neural Responses to Population Behavior Neural Focus Group Predicts Population-Level Media Effects," *Psychological Science, 23*(5), 439-445.

Flanders, T. (1911), *"Speakers Give Sound Advice," Syracuse Post Standard*, March 28.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004), "Intersubject Synchronizzation of Cortical Activity During Natural Vision," *Science, 303*(5664), 1634-1640.

Itti, L., & Koch, C. (2001), "Computational Modelling of Visual Attention," *Nature Reviews Neuroscience, 2*, 194-203.

Klimesch, W. (2012), "Alpha-Band Oscillations, Attention, and Controlled Access to Stored Information," *Trends in Cognitive Science, 16*(12), 606-617.

Lavie, N. (1995), "Perceptual Load as a Necessary Condition for Selective Attention," *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 451-468.

Lavie, N., & Tsal, Y. (1994), "Perceptual Load as a Major Determinant of the Locus of Selection in Visual Attention," *Perception & Psychophysics, 56*(2), 183-197.

Mackay, M., Cerf, M., & Koch, C. (2012), "Evidence for Two Distinct Mechanisms Directing Gaze in Natural Scenes," *Journal of Vision, 12*(4), 1-12.

Ohme, R., Matukin, M., & Pacula-Lesniak, B. (2011), "Biometric Measures for Interactive Advertising Research," *Journal of Interactive Advertising, 11*(2), 60-72.

Pieters, R., Wedel, M., & Batra, R. (2010), "The Stopping Power of Advertising: Measures and Effects of Visual Complexity," *Journal of Marketing, 74*(5), 48-60.

Poggi, J. (2015), "Is the Super Bowl Worth the Price? For these Advertisers, It Might Be," *Advertising Age, 86*.

Poldrack, R. A. (2008), "The Role of fMRI in Cognitive Neuroscience: Where Do We Stand?" *Current Opinion in Neurobiology, 18*(2), 223-227.

Putrevu, S., Tan, J., & Lord, K. R. (2004), "Consumer Responses to Complex Advertisements: The Moderating Role of Need for Cognition, Knowledge, and Gender," *Journal of Current Issues & Research in Advertising, 26*(1), 9-24.

Simion, C., & Shimojo, S. (2006), "Early Interactions Between Orienting, Visual Sampling and Decision Making in Facial Preference," *Vision Research, 46*(20), 3331-3335.

Teixeira, T. S., Wedel, M., & Pieters, R. (2010), "Moment-to-Moment Optimal Branding in TV Commercials: Preventing Avoidance by Pulsing," *Marketing Science, 29*(5), 783-804.

Watters, L. (2017). *"*The History of Super Bowl Commercial Costs," *American Marketing Association*, January 27.