ARTICLE IN PRESS

Journal of Business Research xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Business Research

journal homepage: www.elsevier.com/locate/jbusres

Predicting consumer gaze hits: A simulation model of visual attention to dynamic marketing stimuli

Christopher Rumpf*, Felix Boronczyk, Christoph Breuer

Institute of Sport Economics and Sport Management, German Sport University Cologne, Cologne, Germany

ARTICLE INFO

Keywords:
Eye tracking
Visual attention
Indirect marketing
Simulation model
Artificial neural network

ABSTRACT

The purpose of the present study is to build and test a simulation model for the prediction of gaze hits in the context of dynamic marketing stimuli. Forecasting the attentional effect of dynamic stimuli is of particular interest when it comes to indirect forms of marketing communication such as sponsorship, product placement, or in-game-advertising. Based on large-scale eye tracking data an artificial neural network was trained, providing high predictive accuracy. The model's business applicability is demonstrated with the case of a soccer sponsorship, using media data and color features as model input. The study highlights the value of eye tracking data for the ex-ante valuation of visual communication stimuli which benefits marketing management at the initiation, implementation, and evaluation stages.

1. Introduction

Over the last two decades, eye tracking studies have significantly enhanced the understanding of consumer gaze behavior in response to visual marketing communication (Wedel & Pieters, 2014). For example, there is solid knowledge on the role of visual features in capturing attention (e.g., Pieters & Wedel, 2004), as well as on downstream effects of attention such as brand memory (e.g., Breuer & Rumpf, 2012) and the impact on product choice (e.g., Guerreiro, Rita, & Trigueiros, 2015). Thanks to a substantial and growing body of research, eye tracking has become an important element in the toolbox of marketing science (Romaniuk & Nguyen, 2017).

Despite its indisputable value to better understand the drivers of visual marketing success, the current use of eye tracking in business research leaves room for further enhancements. In the field of commercial business research, eye tracking is best known for its capacity to produce illustrative heat maps, however, they offer very limited predictive validity (Wedel & Pieters, 2014). Additionally, in most cases eye tracking data is utilized for ex-post assessments of marketing activities, while research has not yet exploited its potential to simulate consumer attention as a way to provide ex-ante valuation. Moreover, eye tracking research to date has generated insights into the effects of static stimuli within websites or print magazines, but only limited implications for the effect of dynamic marketing stimuli (Shi, Wedel, & Pieters, 2013; Wedel & Pieters, 2008).

The effects of dynamic marketing stimuli are of particular interest when it comes to indirect forms of marketing communication – such as sponsorship, product placement or in-game-advertising. Over recent years, these so called ‘below-the-line marketing’ instruments have enjoyed major growth relative to traditional forms of marketing communication (IEG, 2017). This ongoing shift from advertising towards more “embedded” and “engaged” ways of marketing communication is likely to be caused by the increasing importance of leisure experiences in modern society and new media behavior in the digital age (Cornwell, 2014).

While embedding brands for instance within a sports competition, video game, or movie represents a promising approach to reach consumers, it also creates new challenges. In professional sponsorship markets, for example, sponsors do not attain placements (e.g., brand logo on player's jersey) on a single basis, but acquire predefined sets of placements – so-called advertising packages. While top-tier sponsors receive full communication rights across all placements, the scope of brand placements for lower tier sponsors is limited (e.g., the brand is only visible on the digital board and on the interview backdrop). Managerial decision-making between such sponsorship options is difficult as ad-hoc data revealing the return on investment is unavailable. Thus, marketing management would benefit from simulation models which generate insights into attention-capturing tactics immediately and before making a decision. Against this backdrop, the study aims at creating a simulation model based on eye tracking data for evidence-based decision making in indirect marketing.

The paper is organized as follows: First, it provides a brief literature

* Corresponding author.
E-mail addresses: c.rumpf@dshs-koeln.de (C. Rumpf), f.boronczyk@dshs-koeln.de (F. Boronczyk), breuer@dshs-koeln.de (C. Breuer).

https://doi.org/10.1016/j.jbusres.2019.03.034
Received 30 July 2018; Received in revised form 18 March 2019; Accepted 20 March 2019
0148-2963/

Please cite this article as: Christopher Rumpf, Felix Boronczyk and Christoph Breuer, Journal of Business Research, https://doi.org/10.1016/j.jbusres.2019.03.034

review of attention to dynamic stimuli, the prediction of marketing outcome, and the use of artificial neural networks in business research. Second, the methodological approach of this study is described, before results are presented on the drivers of visual attention and the simulation model. Then, the application of the simulation model is illustrated with the case of a soccer sponsorship. Finally, the study's contribution, implications, and limitations are critically discussed.

2. Literature review

2.1. Visual attention to dynamic marketing stimuli

Visual attention is typically described as the allocation of an individual's processing capacities to stimuli in their visual field (e.g., Bundesen, Habekost, & Kyllingsbaek, 2005). Information within the "spotlight" of attention is processed more quickly or more efficiently (Brefczynski & DeYoe, 1999). Essentially, objects may receive visual attention through two basic mechanisms: Bottom-up and top-down attention, also referred to as exogenous and endogenous control of attention. Whereas individuals can direct their top-down attention consciously (e.g., Wolfe & Horowitz, 2004), bottom-up attention is allocated automatically and more rapidly (Itti & Koch, 2000). The amount of bottom-up attention is mainly driven by the object's visual saliency (Orquin & Mueller-Loose, 2013). This saliency is determined by factors related to the object, such as its hue, luminance, or the contrast relative to its environment, with more salient objects being attended more quickly or for a longer time (Milosavljevic, Navalpakkam, Koch, & Rangel, 2012; van Zoest, Donk, & Theeuwes, 2004).

Computational models of visual attention assume that the saliency of objects in an individual's visual field is computed in a pre-attentive manner. The results of this pre-attentive processing are encoded in a topographic saliency map, in which more salient objects "pop out" and automatically draw attention towards them (e.g., Itti & Koch, 2000, 2001; Parkhurst, Law, & Niebur, 2002). Such models estimating the saliency of objects in a scene can provide valuable insights for the prediction of attention allocation in static scenes or images. However, a significant number of marketing stimuli, such as TV commercials or sponsor signage in TV broadcasts, are dynamic in nature. In such dynamic marketing stimuli not only the consumer's field of vision, but also the stimuli themselves move, which complicates the study of visual attention.

By analyzing the visually important regions within video clips, Le Meur, Le Callet, and Barba (2007) build a dynamic saliency model which outperforms state of the art models such as the saliency model by Itti and Koch (2000). Furthermore, Mital, Smith, Hill, and Henderson (2011) find that low-level visual features like motion and contrast are most predictive of gaze behavior within dynamic scenes. Interestingly, gaze behavior is significantly more coherent when watching fiction on TV compared to viewing natural scenes (Dorr, Martinetz, Gegenfurtner, & Barth, 2010). Despite the value of these studies to understand gaze behavior in various dynamic contexts, the analytical focus is on primary regions within the scene, and thus, they can barely help to predict consumer attention to brand stimuli which are of secondary importance for the visual processing of a scene.

Only few studies have attempted to examine visual attention to dynamic marketing stimuli (Wedel & Pieters, 2008). The majority of these studies rely on eye tracking data as a widely established measure of visual attention (Venkatraman et al., 2015; Wedel & Pieters, 2014) and focus on consumer attention within TV ads (e.g., Aoki & Itoh, 2000; Janiszewski & Warlop, 1993; Teixeira, Wedel, & Pieters, 2010). However, TV ads represent a direct form of marketing, and only few studies have analyzed visual attention to embedded marketing stimuli like sponsor signage in sports broadcasts (e.g., d'Ydewalle & Tamsin, 1993; Breuer & Rumpf, 2012, 2015) or virtual billboards in video games (e.g., Lee & Faber, 2007).

d'Ydewalle and Tamsin (1993) use eye tracking to measure the incidental processing of sponsor signage during sports broadcasts and find that brands only receive a small share of the sport viewer's attention, namely about 3% of total exposure time. Breuer and Rumpf (2012) also investigate the processing of brand information in a sponsorship context and point out that visual attention is necessary for further cognitive processing of such information. The results of their eye tracking experiment show that visual attention to sponsor signage largely depends on bottom-up factors such as exposure time, size and clutter. A follow-up study by the same authors confirms that – in addition to brand exposure – color and contrast are important determinants for visual attention to sponsor signage (Breuer & Rumpf, 2015). Although these findings reveal important factors in the processing of dynamic marketing stimuli, they still need to be applied to managerial issues in order to adequately address calls for better methods of estimating the return on marketing investments (e.g., Jensen & White, 2018).

2.2. Prediction of marketing communication outcome

One emerging line of marketing research has focused on using experimental data in order to predict effects on the individual level and to forecast marketing outcomes on an aggregate (i.e., target group) level. For instance, several researchers have attempted to use neural response data collected through fMRI studies in order to predict market success (e.g., Berns & Moore, 2012; Falk, Berkman, & Lieberman, 2012; Genevsky, Yoon, & Knutson, 2017), while others have used EEG signals to predict future brand choice (e.g., Barnett & Cerf, 2017; Boksem & Smidts, 2015; Dmochowski et al., 2014; Guixeres et al., 2017; Telpaz, Webb, & Levy, 2015; Venkatraman et al., 2015). Overall, these studies suggest that aggregate behavior can in fact be forecasted reliably based on experimental data (Knutson & Genevsky, 2018).

Zhang, Wedel, and Pieters (2009) study how visual attention data can help improve advertising decisions. The authors match data from eye-tracking tests involving newspaper ads with sales data of the featured product to examine how visual attention for feature advertisements and ad characteristics (e.g., size, color, location) are linked to sales. A Bayesian statistical model based on these data reveals that with greater attention to the feature advertisement sales increase further than through the mere presence of an ad.

In the context of visual attention to dynamic marketing stimuli, Teixeira et al. (2010) examine consumers' attention to brands shown in TV commercials. Based on a dynamic probit model, the authors use attention dispersion metrics to predict TV viewers' advertisement avoidance and use this model to optimize brand exposure within commercials. Suggesting a "pulsing" strategy as the optimal strategy, they further confirm their findings by running a validating experiment.

In an attempt to forecast views of Super Bowl ads on YouTube, Guixeres et al. (2017) create an experimental design and collect a number of biometric data including eye movements, but also heart rate variability and brain response measured through EEG. Using these data, the authors create a prediction model based on an artificial neural network (ANN) which is able to classify and estimate the number of ad views on YouTube with considerable accuracy.

While this application of ANN represents an important first step in the prediction of advertising success in a digital environment, no such prediction models have been developed for the effectiveness of indirect marketing instruments to this point. More specifically, estimations of visual attention to embedded brand stimuli in video content could allow more sophisticated predictions of consumer reactions. Additionally, a simulation model of visual attention can be used to create more reliable, aggregate-level forecasts of indirect marketing effects.

2.3. Artificial neural networks in business research

Decades after artificial neural networks were designed to better understand cognitive processing in the human brain (Bishop, 2006) they have gained increasing popularity in several fields of research as an alternative to multivariate approaches such as regression analysis (DeTienne & DeTienne, 2017). In business research ANNs have been used, for example, to predict consumer response to advertising slogans (Wan-Chen, Santos, & Moutinho, 2016), to forecast the demand for crude oil (Khazem & Mazouz, 2013), or to predict consumer choice for new technologies (Kennedy, Dinh, & Basu, 2016).

Compared to multivariate regression, ANNs are capable to intrinsically deal with multicollinearity, thus they can be useful in addition to regression analysis as a way to enhance the predictive accuracy (Grønholdt & Martensen, 2005; Moore, Beauchamp, Barnes, & Stammerjohan, 2007). Another advantage of ANNs lies in their ability to, per se, include nonlinear relationships between input and output variables (Linder, Geier, & Kölliker, 2004). Since the impact of marketing activities typically reaches a saturation point (Janiszewski, 1993), ANNs have the capacity to enhance the model fit by automatically taking into account such nonlinearities. Further, ANNs consider relevant interaction effects between predictor variables (Grønholdt & Martensen, 2005).

Haykin (2009, p. 2) defines ANN as "a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use". According to this definition, ANN can be understood as an information processing system which models knowledge through "training" and then stores this knowledge in form of coefficients, called "synaptic weights". Within the modelling process, ANNs imitate brain processes by using nonlinear mathematical functions to develop meaningful relationships between input and output variables (Moore et al., 2007).

A basic ANN consists of three layers: The first layer, called the input layer, holds all the exogenous information, for example media exposure parameters. This input layer processes the data to a second, hidden layer, which extracts relevant patterns from the received information. All information which is of importance becomes then transmitted to the output layer, which holds the predicted value for a given performance variable, for example gaze hits. Fig. 1 shows a so-called multilayer perceptron network with four units on the input layer (I1–I4), three units on the hidden layer (H1 – H3), and two units on the output layer (O1 and O2). The units of a preceding layer are linked to all units of a subsequent layer (Grønholdt & Martensen, 2005).

3. Methodology

3.1. Research design

To establish a solid data basis for the prediction of gaze hits within indirect marketing communication, a controlled lab study was designed. Since sport broadcasts typically include dynamic brand stimuli in form of sponsor signage (e.g., on sports apparel or perimeter boards) they represent an ideal research context. In an attempt to consider the various content formats of sport broadcasts, 14 types of sports were recorded, edited and finally used as stimulus material. Prior to collecting eye tracking data, bottom-up factors regarding visible signage, like brand exposure and brand color, were analyzed to serve as predictive variables.

3.2. Stimulus material

The stimulus material comprised video footage representing 14 popular types of spectator sports (i.e., soccer, handball, basketball, athletics, tennis, biathlon, ski alpine, boxing, cycling, motor sports, hockey, swimming, extreme sports, and sailing). The live recordings of each sport broadcast were cut down to a clip of 9 min. Each clip included 3 min of highly arousing scenes (e.g., goals in soccer, passing maneuvers in race sports), 3 min of less exciting scenes, and 3 min of non-sports scenes (e.g., warm-up or half-time interviews). Noteworthy, sponsor signage was clearly visible in each clip.

3.3. Participants

A convenience sample of 315 participants, 63% male and 37% female, was invited to an eye tracking lab. The participants' age ranged between 17 and 64 years, with a mean age of 28.4 years (*SD* = 9.1). The educational level of participants was rather high (91% held a higher education qualification) and income was rather low (68% had < 1500 € per month), both due to a large share of students in the sample. All participants expressed a moderate or high interest towards the assigned sport event. The participation in the experiment was voluntary and people did not receive a monetary incentive in return.

3.4. Measurements

3.4.1. Brand exposure analysis

Each brand appearance with a clear screen visibility of at least 1 s was detected by an automatic image recognition system (Magellan 4.1). The system tracked the exposure time as well as the size, clutter, and position for each brand appearance on screen (see example video frame in Fig. 2). While the size was measured as pixel coverage in relation to the full screen pixels, clutter was based on the number of brands which were detected at the same point in time. The screen position was determined by automatically assigning a score between 1 and 10, with 1 for appearances in the periphery of the screen and 10 for brands appearing in the very center of the screen.

3.4.2. Color detection

Color features were detected by measuring the hexadecimal codes of the colors for each brand appearance on screen. Two different color samples were taken: One sample of the brand logo and one sample of the background (i.e., soccer pitch, basketball court or race track). The hexadecimal codes were transformed into the L*a*b* color model, with 'L' for luminance and 'a' and 'b' for two color opponent dimensions (MacAdam, 1985). On the basis of the L*a*b* model the color contrast ΔE between the background and the logo and was calculated:

$$\Delta E_{back,logo} = \sqrt{(L_{back} - L_{logo})^2 + (a_{back} - a_{logo})^2 + (b_{back} - b_{logo})^2}$$

3.4.3. Measure of visual attention

Eye tracking was used since eye movements are eminent indicators of visual attention (Duchowski, 2007; Holmqvist et al., 2011). To measure gaze hits for the brand logos dynamic areas of interest (AOI) were marked throughout the stimulus material. A table-mounted infrared system (SMI RED) recorded the participants' eye movements with a frequency of 60 Hz. By matching the participant's gaze coordinates with the AOI coordinates in each video frame, the analysis software (SMI BeGaze) was able to count the frequency of gaze hits (with a fixation duration of at least 100 ms). Individual calibration of each participant with nine different calibration points and four validation points on the television screen (42″) was performed to ensure good data quality. In the post-processing of the eye tracking data the sum of 'gaze hits' per brand and participant was calculated. Table 1 provides an overview of variables that were measured in the study.

3.5. Procedure

The setup of the study was standardized for all participants. The procedure was as follows: First, participants received a briefing during which they were told that the study's objective was about emotional reactions to sport telecasts. Subsequent to the eye tracker calibration, the participants were asked to watch one sport clip with a duration of 9 min. In order to support the cover story, the participants were asked

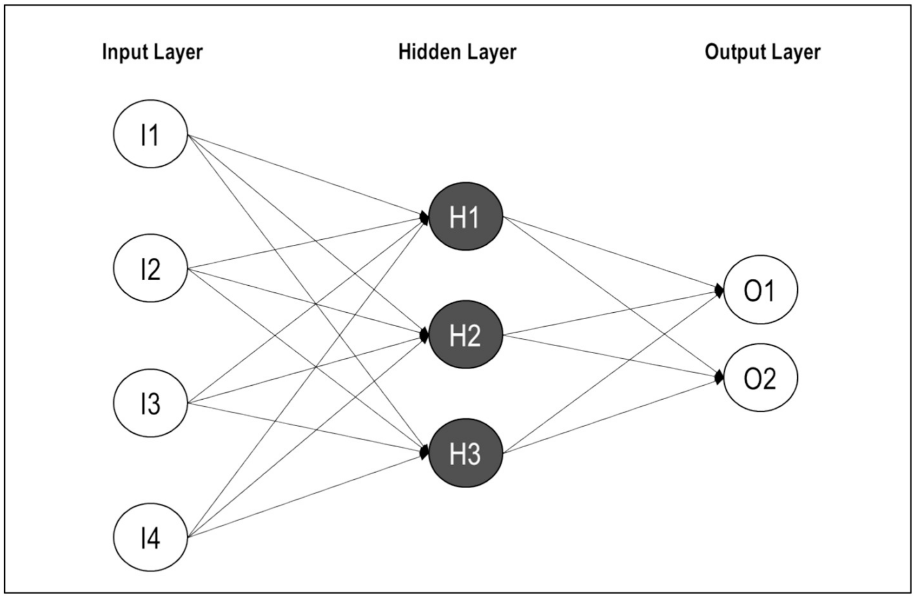

Fig. 1. Model architecture of a basic multilayer perceptron network.

to evaluate their emotional perception on a self-assessment scale after the video presentation.

3.6. Data analysis

The final data set held one case for each participant and brand. Due to the fact that more than one brand appeared in each of the stimulus clips, every visible brand was treated as a new case resulting in $k = 3807$ cases (see Table 2). Based on these count data, a General Linear Mixed Model (GLMM) with Poisson distribution and an ANN were computed. In the GLMM, stimulus clips and participants were treated as random effects since brand visibility differed between stimulus clips and participants were expected to vary in their general interest towards sponsor brands. Both models were estimated with IBM SPSS Statistics 25.0.

4. Results

4.1. Analyzing the drivers of visual attention

General linear mixed regression modelling with random effects nested by participant and stimulus clip is applied to assess the impact

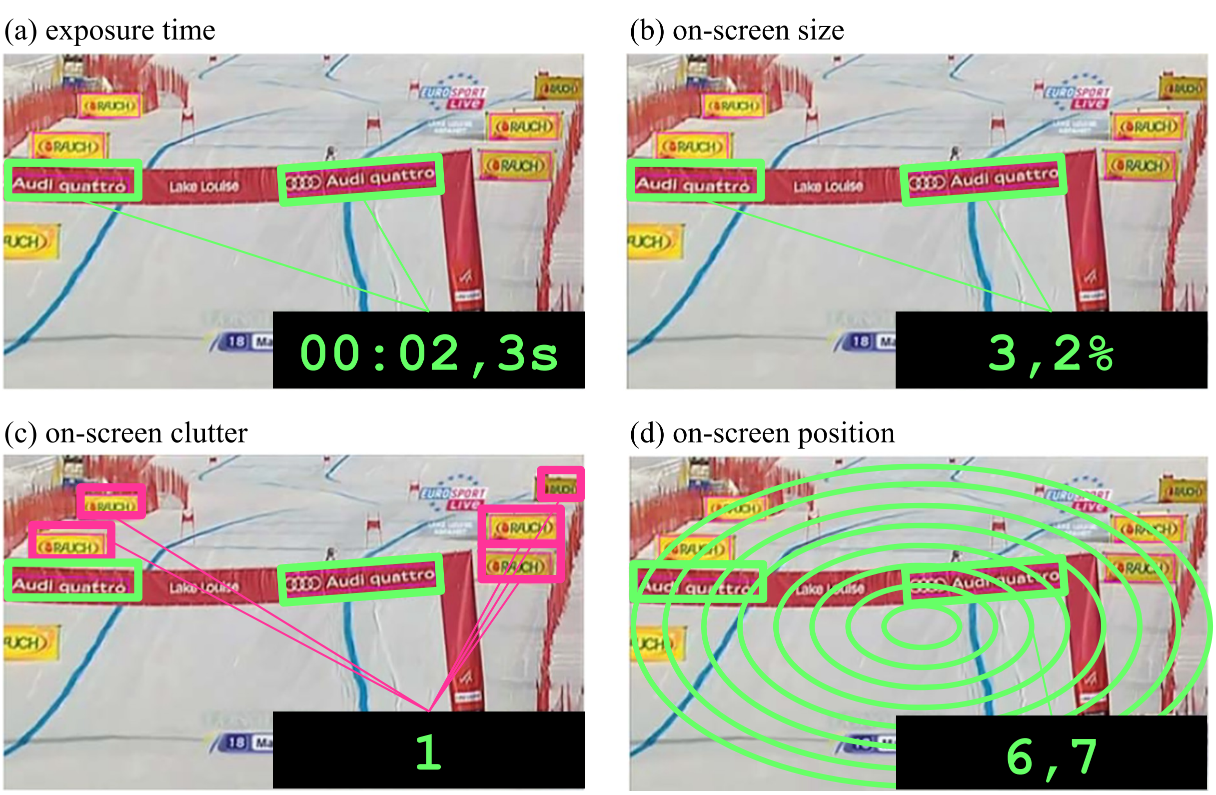

Fig. 2. Example of brand exposure analysis for one video frame.

Table 1
Overview of variables measured in the study.

Method	Variable	Description of measurement
Exposure analysis	Share of visibility [%]	Seconds of brand visibility on screen divided by the duration of the stimulus clip (9 min = 540 s)
	Avg. size on screen [%]	Pixel coverage in relation to full screen pixels (weighted average across brand appearances on screen)
	Avg. clutter on screen	Amount of simultaneously exposed brands (weighted average across brand appearances on screen)
	Avg. position on screen [1 − 10]	Score with 1 for appearances in the periphery and 10 for brands appearing in the very center of the screen (weighted average across brand appearances on screen)
Color detection	L* (luminance)	Luminance channel with 100 = white and 0 = black
	a-value	Red/green channel with + = redder and − = greener
	b-value	Yellow/blue channel with + = yellower and − = bluer
	ΔE (color contrast)	Difference between background and brand logo based on the luminance and chromatic channels (a, b)
Eye-tracking	Gaze hits	Frequency of gaze hits with brands

Table 2
Descriptive results.

Variable	k	Min	Max	Mean	S.D.
Gaze hits	3807	0.00	99.00	4.87	7.797
Share of visibility [%]	3807	0.19	98.70	15.63	17.90
Size on screen [%]	3807	0.10	17.00	1.24	1.40
Clutter on screen	3807	0.00	7.00	2.73	1.55
Position on screen [1–10]	3807	1.00	8.90	4.29	1.44
L* (luminance)	3807	0.59	99.53	54.73	27.47
+ a* (reddish)	3807	0.00	80.50	11.78	19.00
+ b* (yellowish)	3807	0.00	90.97	13.80	21.92
− a* (greenish)	3807	0.00	53.43	3.45	8.11
− b* (bluish)	3807	0.00	78.09	6.82	13.12
ΔE (color contrast)	3807	3.64	101.85	51.32	20.29

on visual attention. All in all, 10 variables are entered into the model, reflecting brand exposure and color features, while 'gaze hits' serves as the outcome variable. The F-statistic shows a high level of significance ($F(10, 3421) = 1022.239$; $p < .001$). Table 3 summarizes the GLMM results.

Among the brand exposure variables, the 'share of visibility' ($b = 0.293$, $p < .001$), 'size on screen' ($b = 0.566$, $p < .001$), and 'position on screen' ($b = 0.274$, $p < .001$) facilitate the capture of 'gaze hits'. As a negative driver, the 'clutter on screen' reduces the degree of visual attention ($b = -0.572$, $p < .001$). Besides brand exposure, the color variables significantly influence the allocation of attention. Based on the L*a*b* color space, luminance (L*), and two color channels (+a*/−a* and +b*/−b*) were assessed for each brand logo. The model shows that the luminance of sponsor colors does not significantly influence the degree of attention, while more reddish (+a*, $b = 0.004$, $p = .020$) and greenish (−a*, $b = 0.009$, $p = .031$) logos receive more visual attention than bluish logos (−b*, $b = -0.027$, $p < .001$). The effect of more yellowish logos (+b*) on 'gaze hits' turns out to be insignificant. Independent from the effect of hue, the color contrast (ΔE) between the brand logo and its surrounding has a positive effect on 'gaze hits' ($b = 0.012$, $p < .001$).

Table 3
Generalized linear mixed model of visual attention.

Parameter	b	SE	*p*	95% CI	
				LL	*UL*
Intercept	−0.914	0.190	< .001	−1.286	−0.541
Share of visibility [%]	0.293	0.003	< .001	0.287	0.299
Size on screen [%]	0.566	0.025	< .001	0.517	0.614
Clutter on screen	−0.572	0.020	< .001	−0.612	−0.532
Position on screen [0 − 10]	0.274	0.018	< .001	0.240	0.308
L* (luminance)	< 0.001	0.001	.881	−0.002	0.002
+ a* (reddish)	0.004	0.002	.020	0.001	0.007
+ b* (yellowish)	−0.001	0.002	.527	−0.005	0.002
− a* (greenish)	0.009	0.004	.031	0.001	0.017
− b* (bluish)	−0.027	0.001	< .001	−0.029	−0.024
ΔE (color contrast)	0.012	0.002	< .001	0.008	0.015
F (10, 3421)	1022.239				
p	*p* < .001				

Note. Dependent variable: Gaze hits.
Probability distribution: Poisson.
Random effect: video clips*subjects.
Cases included in the model: = 3807.

4.2. Predicting visual attention

The GLMM tests for the effect size and the significance of the prediction variables, while the ANN is applied to maximize the predictive accuracy. Given the issue of overtraining (e.g., Moore et al., 2007), the data ($k = 3807$) are randomly split into two sub-samples: The first sub-sample, ($k = 2923$), is used to train the ANN; the second sub-sample, namely the test sample ($k = 884$), is an independent class of data sets used to track errors during the ANN development process to prevent overtraining.

Literature provides contradicting recommendations on the selection of the appropriate network architecture. Since there is no well-developed theory of network architecture optimization (Haykin, 2009), we have followed the common approach to add nodes on each hidden layer in an effort to reach the optimal model performance in terms of minimized error. By following this trial-and-error process, the final network architecture with ten nodes in both hidden layer has been identified (Fig. 3).

Based on a multilayer perceptron (MLP) with a backpropagation algorithm, the ANN is trained to predict 'gaze hits' until the squared error within the total sample (both training and test sample) is optimized. In an effort to analyze the importance of the predictive variables, a sensitivity analysis is conducted which determines the change in the output value based on a gradual increase or decrease of the input values (Therón & Paz, 2006). Fig. 4 displays the relative importance of all predictive variables with the strongest predictor ('share of visibility') set to 100%.

In a second step, the prediction accuracy is to be analyzed by comparing the ANN to the GLMM. Fig. 5 provides a comparison of the ANN and the GLMM in form of scatterplots which show that the model fit is substantially higher in the ANN ($R^2 = 0.822$) compared to the GLMM ($R^2 = 0.462$).

However, model fit is a poor indicator of predictive accuracy since the same data is used for building and testing the model. To assess model generalizability, a 5 × 2 iterated cross-validation is run with stratified random sampling. Assuming that the viewing behavior differs due to individual characteristics and video-related differences, the cross-validation folds were stratified by (1) subject and (2) clip. In both instances, the total sample is randomly split into two independent subsamples (fold A + B), whereby one half of the data is used for building a model and the other half is used for testing its predictive accuracy as indicated by the mean squared error (MSE). Table 4 displays the cross-validation results for 5 × 2 iterations based on randomized subject sampling.

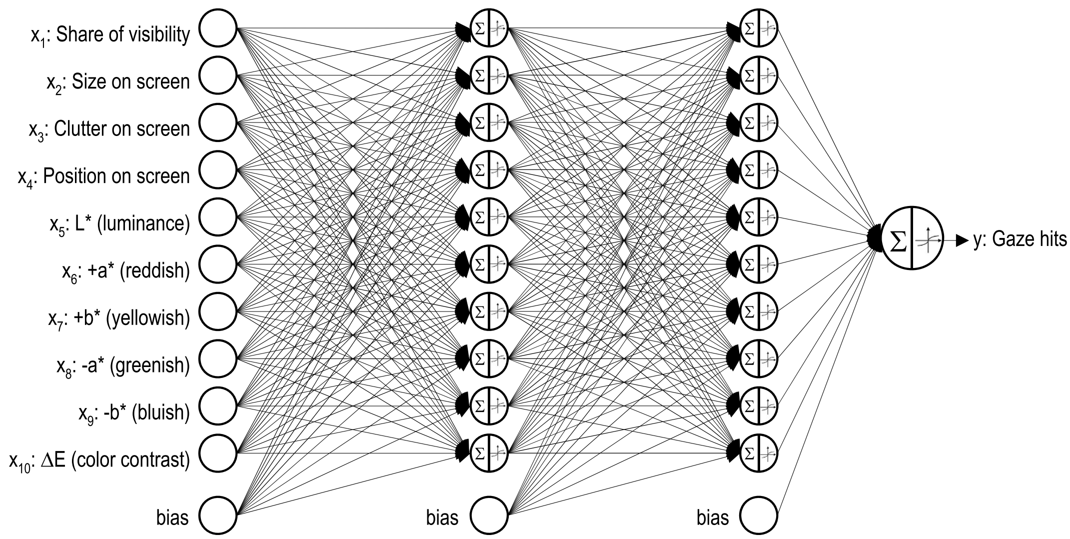

Fig. 3. ANN of visual attention.

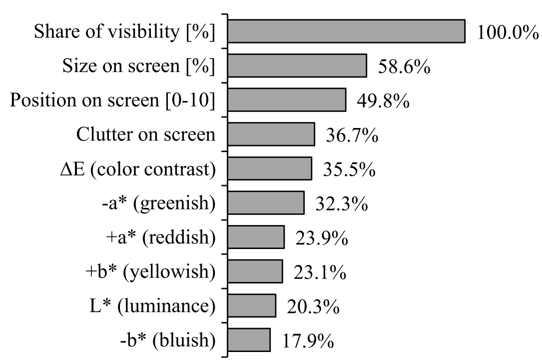

Fig. 4. Sensitivity analysis.
Note. Importance of each predictor in determining the ANN; Combined training and test sample (k = 3807).

The ANN creates lower MSE for all iterations with highly significant t-values in nine out of ten cases. The procedure is replicated for the stratified clip sampling (see Table 5), showing that the performance of the ANN is not superior to the GLMM when taking into account the differences between stimulus clips. Apparently, the backpropagation algorithm of the ANN is stronger in discovering regularities between the viewing characteristics of the subjects than systematic patterns between different kinds of sport broadcasts. Given that the ANN clearly outperforms the GLMM in the cross-subject validation and leads to comparable accuracy with regard to the clip-based validation approach, the ANN is identified as the more powerful solution for the prediction of gaze hits in the context of dynamic marketing stimuli.

5. Application of the simulation model

After building and testing the simulation model, its applicability is

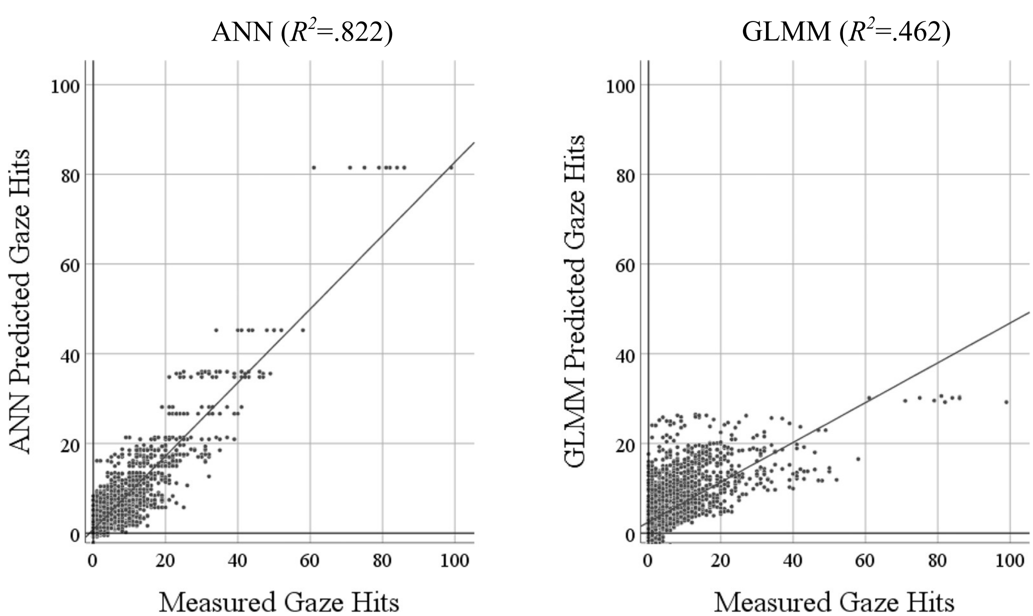

Fig. 5. Comparison by scatterplots.

Table 4
5 × 2 iterated cross validation based on subject sampling.

Iteration	Fold size		A → B			B → A		
	A	B	ANN	GLMM	*t*	ANN	GLMM	*t*
1	1840	1967	38.2	42.1	−0.466	19.5	35.1	−4.166***
2	1976	1831	15.5	36.2	−5.540***	17.4	40.9	−4.806***
3	1904	1903	13.8	37.2	−5.350***	21.1	39.6	−4.195***
4	1876	1931	20.4	43.4	−4.500***	14.0	33.9	−5.650***
5	1827	1980	15.7	43.2	−5.372***	15.9	35.4	−5.680***

Note. Values indicate MSE (Mean Squared Error) for holdout sample.
ANN = Artificial Neural Network; GLMM = Generalized Linear Mixed Model.
A → B: Model built on subsample A, validation on holdout sample B.
B → A: Model built on subsample B, validation on holdout sample A.
$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$ (two-tailed).

Table 5
5 × 2 iterated cross validation based on clip sampling.

Iteration	Fold size		A → B			B → A		
	A	B	ANN	GLMM	*t*	ANN	GLMM	*t*
1	1600	2207	59.9	51.7	1.022	29.8	26.3	1.244
2	2215	1592	29.2	39.6	−3.524***	64.0	56.6	0.753
3	1320	2487	44.1	36.7	1.674	68.2	46.8	1.765
4	1918	1889	57.8	40.6	2.043*	38.0	40.0	−0.420
5	1620	2187	52.8	43.9	0.979	44.9	42.0	0.702

to be presented. For this purpose, media exposure data were provided by a leading media agency covering the case of soccer sponsorship in the German Bundesliga. In this case, brand logos are visible at different placements, for instance, around the pitch (e.g., digital board), next to the goals (e.g., cam carpet), or on the players' team kit (e.g., jersey).

The case data allows predicting 'gaze hits' for each placement. For example, the brand logo on the jersey reached a visibility share of 4.7% (in relation to the total broadcasting time) with an average on-screen size of 1.6% (in relation to the full screen). Moreover, the case data reveals that – on average – 0.9 other brands were exposed simultaneously (i.e., clutter on screen) and that the logo on the jersey appeared in a mid-central screen position (5.2 on a 1–10 scale). In terms of colors, this example is based on a red brand placed on a white jersey which comes with a considerable color contrast (ΔE = 88.1). By feeding the simulation model with the aforementioned input data, a prediction value of 0.98 'gaze hits per minute' for the brand on the jersey is derived.

In professional sponsorship markets, however, brands do not choose between single logo placements (e.g., on the jersey), but attain so-called advertising packages. In this regard, the simulation model can assist managers in their decision making by predicting the degree of attention to be expected by different advertising packages. To be more specific, a tier-1 sponsorship provides high exposure levels compared to a tier-2 sponsorship with less media exposure. At the same time, the range of different color designs is taken into account. Based on the real-world media exposure data the simulation model predicts the 'gaze hits per minute' as a measure of attention capture. Fig. 6 provides boxplots for both sponsorship options.

Not very surprisingly, the boxplots show a higher predicted outcome for the tier-I-sponsorship, whereas the variability is considerably larger due to the impact of sponsor logo colors. The importance of the color design can be illustrated by the example of a blue logo (brand A) vs. an orange logo (brand B). Brand A strives for the tier-I-sponsorship which guarantees logo visibility at the most prominent locations in the stadium and on the players' apparel (Fig. 7).

Brand B plans for a tier-II-sponsorship with a reduced set of placements in the stadium. It is noteworthy that logo visibility on the digital board is not continuous, but changes during the game with more frequent appearances of top-tier sponsors (Fig. 8).

Taking these settings into account the simulation model predicts slightly more gaze hits for brand B (pred. 'gaze hits per minute' = 0.92) compared with brand A (0.90), despite more prominent placements. This example stresses the relevance of assessing the degree of attention capture instead of relying on media exposure figures as success indicator in indirect marketing communication. Additionally, the amount of attention a brand receives is not only dependent on the platform characteristics (e.g., the visibility of placements), but also on its visual features.

6. Discussion

This paper contributes to the body of knowledge as it provides an applicable simulation model for the ex-ante valuation of indirect marketing communication. The model was trained based on eye tracking data across 14 different types of sport in an effort to enhance its generalizability.

Compared to the mixed regression model, the ANN provides very accurate and robust predictions regarding different sets of consumers, that is, interindividual differences. When taking into account the broad range of sport venues and telecasts, the effect of overtraining cannot be excluded. The problem of overtraining can occur if the ANN "memorizes" particular features which are present in the training data, but absent in the separate settings that are to be predicted (Haykin, 2009). Given that brand exposure within sport broadcasts can differ enormously depending on the sport setting (e.g., basketball on an indoor court vs. motor sports on an outdoor circuit) overtraining could have occurred. If so, the context of the training data might need to be narrowed down (e.g., only consider one particular type of sport) in the model's further development.

The model application using a sponsorship case in soccer demonstrates how the simulation model considers the effect of low-level visual features to predict human gaze behavior. While previous work by Itti and Koch (2000, 2001) relied on capturing the visual properties of static images to construct saliency, the simulation model presented here predicts the human response to such bottom-up characteristics based on actual gaze behavior observed in eye-tracking tests.

In vision science, several studies have extended the understanding of gaze behavior in dynamic visual scenes (e.g., Dorr et al., 2010; Le Meur et al., 2007; Mital et al., 2011), but their focus has been on the primary stimulus. In the context of sponsorship, in-game advertising, or product placement, brand stimuli appear embedded into the visual scene, for example on jerseys, virtual billboards, or branded products. Since the viewer's main interest is focused on primary stimuli (e.g., athletes, game characters or actors), the attention-grabbing features are likely to differ. Only few studies have examined visual attention to brand stimuli in dynamic media (Wedel & Pieters, 2008), and to the authors' best knowledge, this simulation model represents the first

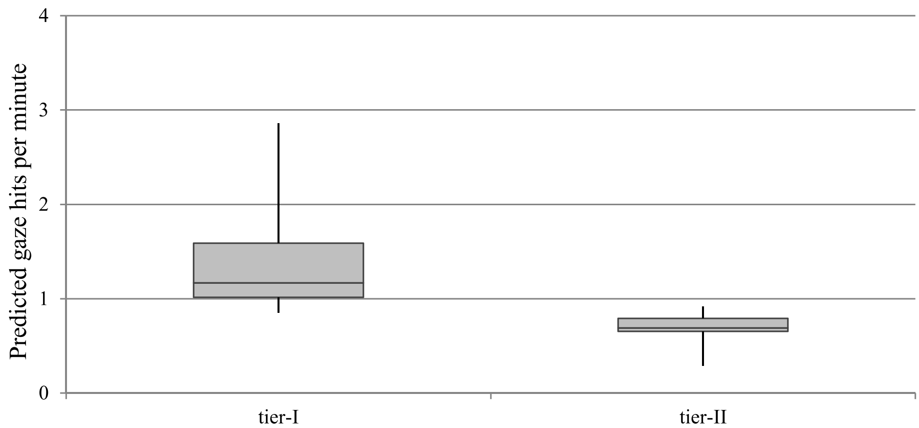

Fig. 6. Variation between sponsorship tiers.
Note. Tier-I-sponsorship: Share of visibility 30.5%, Avg. size on screen 2.5%, Avg. clutter on screen 2.8, Avg. position on screen 3.9; Tier-II-sponsorship: Share of visibility 15.4%, Avg. size on screen 2.4%, Avg. clutter on screen 3.5, Avg. position on screen 3.7.

attempt to forecast visual attention in the context of indirect marketing communication.

The model benefits marketing researchers and professionals as it allows predicting the attentional outcome of visual communication activities. Thus, different activities can be systematically compared with regard to their effectiveness (number of gaze hits) and efficiency (gaze hits in relation to marketing spending). In this way, the model supports evidence-based decision making and helps to enhance the professional management of indirect marketing. In the remainder, the simulation model will be discussed concerning the initiation, implementation, and evaluation of marketing communication activities:

1) Initiation: Marketing management compares the value of a marketing communication opportunity to alternative options.
2) Implementation: Marketing management exploits the communication rights in an effort to maximize effectiveness.
3) Evaluation: Marketing management realigns elements of the communication activity.

During the initiation phase, marketing management is faced with the challenge to identify the most effective and efficient opportunity from a range of alternative options. Given a good fit with the brand values and the key target group, managers will choose the option which will generate the best cost-benefit ratio. While the costs of a marketing activity can be calculated based on property right fees and prices for media services, the return of an indirect marketing activity is often difficult to foresee. To cope with this uncertainty, the simulation model allows comparing ex-ante the degree of consumer attention to the brand message as a key variable in consumer behavior (Wedel & Pieters, 2014).

The simulation model relies on input data concerning media exposure and brand color features. In the context of sponsorship-linked communication, media exposure data is mostly available through property right holders (i.e., event organizer, sport marketing agency), while brand colors are part of the corporate design and thus easy to identify. Based on the input data, the model calculates the gaze hits per minute for one consumer. By multiplying this prediction value by the broadcasting time and the audience reach of the event, total attention to the brand can be forecasted on an aggregate level, which allows a holistic appraisal of the marketing success. Finally, this value can be put in relation to the overall marketing efforts to assess the efficiency.

During the implementation of marketing activities it is of major importance to monitor marketing performance (Olson & Thjømøe, 2009; O'Reilly & Madill, 2009). In the context of indirect marketing, most brand managers rely on media exposure parameters, even though these figures cannot reflect communication effectiveness. Instead of maximizing the level of brand exposure, managers should strive for the highest possible degree of attention (Breuer & Rumpf, 2012). By predicting the gaze hits per minute, the simulation model provides a performance indicator that serves as a valid measure of consumer attention. In contrast to media exposure parameters the prediction of gaze hits is sensitive to design features such as the color of the brand logo. Further, gaze hits as a single measure integrates the information of several exposure variables (duration, size, clutter, position) and thus reduces the complexity of drawing conclusions.

Finally, the simulation model provides a solid basis for decision

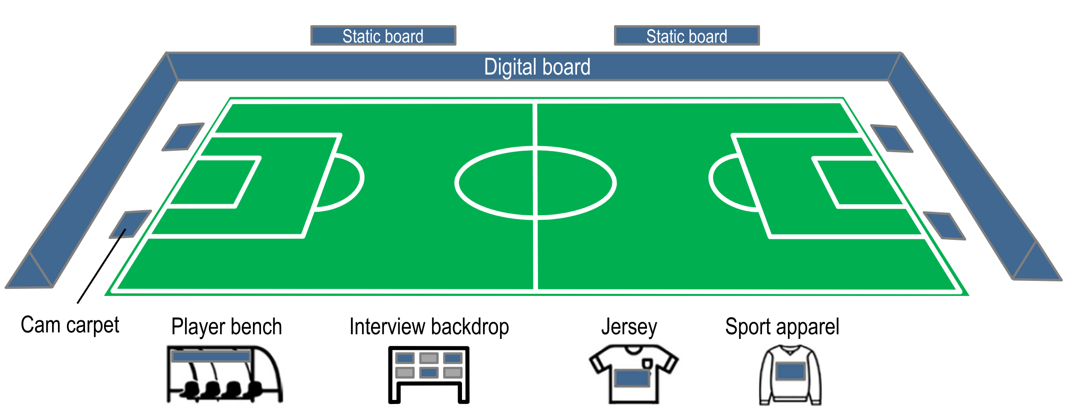

Fig. 7. Tier-I logo placements of brand A.
Note. Share of visibility 30.5%, Avg. size on screen 2.5%, Avg. clutter on screen 2.8, Avg. position on screen 3.9; L* 42.81, a* −1.47, b* −26.77, ΔE 74.61.

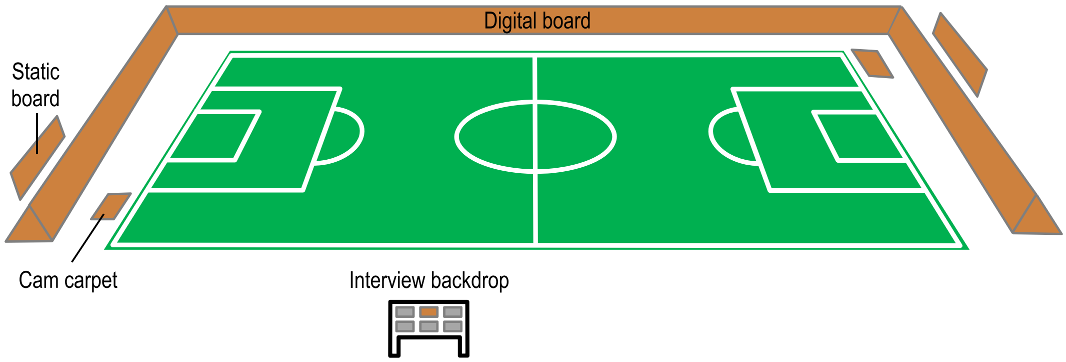

Fig. 8. Tier-II logo placements of brand B.
Note. Share of visibility 15.4%, Avg. size on screen 2.4%, Avg. clutter on screen 3.5, Avg. position on screen 3.7; L* 60.81, a* 23.51, b* 47.37, ΔE 58.34.

making when it comes to realigning the communication elements. The brand manager might consider upgrading to a higher placement package or employing different colors. Moreover, the conditions under which the brand is exposed to consumers might change in terms of clutter (more brands visible simultaneously) or different background colors. To analyze in how far changes affect the value of a marketing activity, the simulation model can immediately calculate the effects in terms of attention. This sets a fair basis for renegotiations, for example to justify a price premium.

7. Limitations and further research

After discussing the study's contribution and implications some shortcomings need to be critically reflected. First, the simulation model might be criticized for being based on data from a lab study in which short sport clips were presented. It can be assumed that the participants' emotional experience was less intensive compared to watching live sport broadcasts. Since the emotional experience of watching sports might have an influence on the allocation of attention, the study results should be replicated in a live sports context.

Second, the simulation model does not account for animated stimuli. Taking into account recent findings from the field of sponsorship-linked communication (Breuer & Rumpf, 2015) and online banner advertisement (Hamborg, Bruns, Ollermann, & Kaspar, 2012), the model should be extended by including animation effects.

Third, the simulation model is based on a multilayer perceptron with a squared-error optimization criterion which is the common approach in ANN modelling (e.g., Bishop, 2006; Haykin, 2009). However, given that the outcome variable 'gaze hits' is measured as count data, Fallah et al. (2009) suggest a hybridization of ANN and general linear modelling in an effort to combine the likelihood measure with the advantages of ANN in handling nonlinearity and noise. Even though it would have been interesting to see if such a model extension impacts the predictive accuracy, this was beyond the scope of this study but should be taken into account in future research.

Finally, the model's strength lies in predicting the attentional outcome of marketing activities in sports, while the impact on brand preference or product choice is not yet considered. Follow-up research is needed to examine the degree to which gaze hits cause downstream effects in an attempt to provide a more complete forecast of marketing success.

Declarations of interest

None.

References

Aoki, H., & Itoh, K. (2000). Analysis of cognitive attitudes to commercial films on basis of eye tracking data. *Proceedings of the human factors and ergonomics society annual meeting. Vol. 44. Proceedings of the human factors and ergonomics society annual meeting* (pp. 38–41). Los Angeles, CA: SAGE Publications (No. 1).

Barnett, S. B., & Cerf, M. (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.

Berns, G. S., & Moore, S. E. (2012). A neural predictor of cultural popularity. *Journal of Consumer Psychology, 22*, 154–160.

Bishop, C. M. (2006). *Pattern recognition and machine learning.* New York: Springer.

Boksem, M. A. S., & Smidts, A. (2015). Brain responses to movie trailers predict individual preferences for movies and their population-wide commercial success. *Journal of Marketing Research, 52*, 482–492.

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience, 2*(4), 370.

Breuer, C., & Rumpf, C. (2012). The viewer's reception and processing of sponsorship information in sport telecasts. *Journal of Sport Management, 26*(6), 521–531.

Breuer, C., & Rumpf, C. (2015). The impact of color and animation on sports viewers' attention to televised sponsorship signage. *Journal of Sport Management, 29*(2), 170–183.

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review, 112*(2), 291–328.

Cornwell, T. B. (2014). *Sponsorship in marketing: Effective communications through sports, arts and events.* New York: Routledge.

DeTienne, K. B., & DeTienne, D. H. (2017). Neural networks in strategic marketing: Exploring the possibilities. *Journal of Strategic Marketing, 25*(4), 289–300. https://doi.org/10.1080/0965254X.2015.1076881.

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, (Article number: 4567).

Dorr, M., Martinetz, T., Gegenfurtner, K., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision, 10*(10), 1–17.

Duchowski, A. (2007). *Eye tracking methodology, theory and practice* (2nd ed.). London: Springer.

d'Ydewalle, G., & Tamsin, F. (1993). On the visual processing and memory of incidental information: Advertising panels in soccer games. *Visual search 2: Proceedings of the 2nd international conference on visual search* (pp. 401–408). London: Taylor & Francis.

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*, 439–445.

Fallah, N., Gu, H., Mohammad, K., Seyyedsalehi, S., Nourijelyani, K., & Reza Eshraghian, M. (2009). Nonlinear Poisson regression using neural networks: A simulation study. *Neural Computing and Applications, 18*(8), 939–943. https://doi.org/10.1007/s00521-009-0277-8.

Genevsky, A., Yoon, C., & Knutson, B. (2017). When brain beats behavior: Neuroforecasting crowdfunding outcomes. *The Journal of Neuroscience, 37*, 8625–8634.

Grønholdt, L., & Martensen, A. (2005). Analysing customer satisfaction data: A comparison of regression and artificial neural networks. *International Journal of Market Research, 47*(2), 121–130.

Guerreiro, J., Rita, P., & Trigueiros, D. (2015). Attention, emotions and cause-related marketing effectiveness. *European Journal of Marketing, 49*(11/12), 1728–1750. https://doi.org/10.1108/EJM-09-2014-0543.

Guixeres, J., Bigné, E., Ausín Azofra, J. M., Alcañiz Raya, M., Colomer Granero, A., Fuentes Hurtado, F., & Naranjo Ornedo, V. (2017). Consumer neuroscience-based metrics predict recall, liking and viewing rates in online advertising. *Frontiers in Psychology, 8*, 1808. https://doi.org/10.3389/fpsyg.2017.01808.

Hamborg, K.-C., Bruns, M., Ollermann, F., & Kaspar, K. (2012). The effect of banner animation on fixation behavior and recall performance in search tasks. *Computers in Human Behavior, 28*, 576–582. https://doi.org/10.1016/j.chb.2011.11.003.

Haykin, S. (2009). *Neural networks and learning machines* (3rd ed.). Upper Saddle River: Pearson Prentice Hall.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford: Oxford University Press.

International Events Group (IEG) (2017). Sponsorship spending forecast: Continued growth around the world. 9Retrieved from the IEG website on 2018/01/26) http://www.sponsorship.com/IEGSR/2017/01/04/Sponsorship-Spending-Forecast–Continued-Growth-Ar.aspx.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506. https://doi.org/10.1016/S0042-989(99)00163-7.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203.

Janiszewski, C. (1993). Preattentive mere exposure effects. *Journal of Consumer Research, 20*(3), 376–392.

Janiszewski, C., & Warlop, L. (1993). The influence of classical conditioning procedures on subsequent attention to the conditioned brand. *Journal of Consumer Research, 20*(2), 171–189.

Jensen, J. A., & White, D. W. (2018). Trends in sport sponsorship evaluation and measurement: Insights from the industry. *International Journal of Sports Marketing and Sponsorship, 19*(1), 2–10.

Kennedy, M., Dinh, V., & Basu, B. (2016). Analysis of consumer choice for low-carbon technologies by using neural networks. *Journal of Cleaner Production, 112*(4), 3402–3412. https://doi.org/10.1016/j.jclepro.2015.10.035.

Khazem, H., & Mazouz, A. (2013). Forecasting the price of crude oil using artificial neural networks. *International Journal of Business, Marketing, & Decision Science, 6*(1), 119–135.

Knutson, B., & Genevsky, A. (2018). Neuroforecasting aggregate choice. *Current Directions in Psychological Science, 27*(2), 110–115.

Le Meur, O., Le Callet, P., & Barba, D. (2007). Predicting visual fixations on video based on low-level visual features. *Vision Research, 47*(19), 2483–2498.

Lee, M., & Faber, R. J. (2007). Effects of product placement in online games on brand memory. *Journal of Advertising, 36*, 75–90.

Linder, R., Geier, J., & Kölliker, M. (2004). Artificial neural networks, classification trees and regression: Which method for which customer base? *Journal of Database Marketing & Customer Strategy Management, 11*(4), 344–356.

MacAdam, D. L. (1985). *Color measurement: Theme and variations* (2nd ed.). New York: Springer.

Milosavljevic, M., Navalpakkam, V., Koch, C., & Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology, 22*(1), 67–74.

Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation, 3*(1), 5–24. https://doi.org/10.1007/s12559-010-9074-z.

Moore, R. S., Beauchamp, M. B., Barnes, C. Z., & Stammerjohan, C. A. (2007). Artificial neural networks: An introduction and application in a consumer behavior setting. *Marketing Management Journal, 17*(1), 27–38.

Olson, E. L., & Thjømøe, H. M. (2009). Sponsorship effect metric: Assessing the financial value of sponsoring by comparisons to television advertising. *Journal of the Academy of Marketing Science, 37*(4), 504–515. https://doi.org/10.1007/s11747-009-0147-z.

O'Reilly, N., & Madill, J. (2009). Methods and metrics in sponsorship evaluation. *Journal of Sponsorship, 2*, 215–230.

Orquin, J. L., & Mueller-Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica, 144*(1), 190–206.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research, 42*(1), 107–123.

Pieters, R., & Wedel, M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing, 68*, 36–50. https://doi.org/10.1509/jmkg.68.2.36.27794.

Romaniuk, J., & Nguyen, C. (2017). Is consumer psychology research ready for today's attention economy? *Journal of Marketing Management, 33*(11 − 12), 909–916. https://doi.org/10.1080/0267257X.2017.1305706.

Shi, S. W., Wedel, M., & Pieters, R. (2013). Information acquisition during online decision making: A model-based exploration using eye-tracking data. *Marketing Science, 59*(5), 1009–1026. https://doi.org/10.1287/mnsc.1120.1625.

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.

Telpaz, A., Webb, R., & Levy, D. J. (2015). Using EEG to predict consumers' future choices. *Journal of Marketing Research, 52*(4), 511–529.

Therón, R., & Paz, J. F. (2006). Visual sensitivity analysis for artificial neural net-works. In E. Corchado, H. Yin, V. Botti, & C. Fyfe (Eds.). *Intelligent data engi-neering and automated learning - IDEAL 2006* (pp. 191–198). Berlin: Springer.

van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 746–759.

Venkatraman, V., Dimoka, A., Pavlou, P. A., Vo, K., Hampton, W., Bollinger, B., ... Winer, R. (2015). Predicting advertising success beyond traditional measures: New insights from neurophysiological methods and market response modeling. *Journal of Marketing Research, 52*, 436–452.

Wan-Chen, W., Santos, M. M., & Moutinho, L. (2016). Modelling consumer responses to advertising slogans through artificial neural networks. *International Journal of Business & Economics, 15*(2), 89–116.

Wedel, M., & Pieters, R. (2008). A review of eye tracking research in marketing. In N. K. Malhotra (Ed.). *Review of marketing research* (pp. 123–147). Emerald Group Publishing.

Wedel, M., & Pieters, R. (2014). Looking at vision: Eye/face/head tracking of consumers for improved marketing decisions. In L. Moutinho, E. Bigné, & A. K. Manrai (Eds.). *Routledge companion on future of marketing* (pp. 177–189). Abingdon: Routledge.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5*(6), 495–501.

Zhang, J., Wedel, M., & Pieters, R. (2009). Sales effects of attention to feature advertisements: A Bayesian mediation analysis. *Journal of Marketing Research, 46*(5) (696-681).