

RSC, Number 5, Issue 2, May 2013, pp. 98-115.

Using the method of observation in testing media advertising

Andrej Kovačič

School of Advanced Social Studies in Nova Gorica, Slovenia

andrej.kovacic@ceos.si

Abstract: *Traditional self-report measures are significantly influenced by cognitive and research bias. Therefore in a quest for alternative methodologies in measuring advertising affects “stopping power” was introduced and measured on 9 different locations. The results of the two hypotheses showed first that this method of measuring is reliable with Krippendorff’s alpha (α) larger than 90 % between the two evaluators. Thus less than 10 % of judgements could be a coincidence, which makes this method highly reliable. Secondly the suggested method was able to identify the differences between two ads with the same content printed in different techniques on all 9 locations.*

Keywords: advertising, research methodology, observation

INTRODUCTION

Advertising expenditure is constantly increasing and the search for an efficient method of testing advertisements has never been as important as in the 21st century. Till now we have had to agree with Poels and Dewitte (2006) that "... advertising literature is not straightforward on what instrument provides the most valid measurement" and a construct of how to measure advertising effectiveness has no "perfect" solution. Measuring advertising is getting more and more complicated as advertisers use increasing levels of emotional content. On this notion Young (2004) claims the issue is not "whether emotions in advertising matter, the problem is how to measure them."

This article aims to test an alternative method for measuring advertising effects, which is based on a stopping power. We first describe the limitations of traditional self-report measures. Then the tested method of observation is described in details. After identifying the research objective and research method we present the results of our study and a conclusion with further research suggestions.

Limitations of traditional self-report measures - reasons to use alternative methods

As presented in Heath and Feldwick (2007) practitioners have adopted an empirical quantitative approach to the measurement of emotions relying mainly on self-reports. Verbal self-report measures are quick, inexpensive and usually do not require special expert knowledge. Typically many items,

like recall and cognitive associations (unipolar or bipolar), are used because of their simplicity and the ability to measure how consumers store product information in memory. Verbal self-report measures, however, do not always provide the understanding of consumer responses when advertising is not predominantly verbal in nature (Zambardino and Goodfellow, 2007; Shapiro, MacInnis and Heckler, 1998; Vakratsas and Ambler, 1999). The three major limitations are: first the respondent has to process information and describe the feeling; second the respondent is influenced by a researcher posing the questions and third it is difficult to measure long-term effects of advertising.

Cognitive bias

A huge critic Hall (2001) showed evidence that with traditional verbal self-report survey “no matter how ingenious we are with projective techniques and semantic scales, cognitive bias is inevitable.” Hall (2004) also showed evidence that with traditional verbal self-report survey “until we can accurately measure the non-cognitive and emotional dimensions of advertising response, we are measuring very little of relevance.” Thus Zambardino and Goodfellow (2007) claim that part of the current increased popularity of non self-report techniques comes from the fact that they do not rely on our conscious verbal recollections as it is not necessary to be aware of advertising communication to have been influenced by it. Similarly Erevelles (1998) reveals another weakness of these methods, namely that both memory and cognitive activity are required to measure the affect. Thus the main reason in favour of an alternative measurement is

that no cognitive activity is required to measure affect, and consequently, there is less “bias”.

Research bias

Braun and Zaltman (2006) recognized the important role of research bias in academic research. They claim that because participants are “aware that the researcher is interested in assessing their attitude changes based on the advertising exposure, some may overestimate the impact of advertising by indicating more favourable attitudes; others may do the reverse and underestimate, because they do not want to believe that the advertising had an impact on their beliefs.” In addition Braun’s and Zaltman’s research shows that respondents tend to give answers that appeal the interviewer.

Long-term effects in advertising

If we define the goal of advertising as the creation of lasting memories it is important to be able to detect its long-term effects as well. Young suggests (Young, 2009): if an ad “does not leave some kind of lasting trace behind in the long-term memory of a consumer, it is difficult to argue that it had any kind of effect.” Also Hall (2004) points out that the limitations of the majority of copy-testing measures are that they tend to “focus on the least important part of advertising: short-term sales effects.”

Suggested method of observation

To limit the above described limitations of traditional self-report measures we have tested a method of observation to measure advertising effects. Techniques based on observation are nothing new in advertising. As pointed out by Berg (1989) it is essential to assure as realistic as possible environment for measurement. Similarly Del Vecchio (1988) suggests method of observation as an addition to classical research methods and as a stand alone method when other methods fail to produce quality data. As suggested by Cradit, Tashchian and Hofacker (1994) the overall reliability of this method is improved by repeated measurements.

The method of observation is suggested by Walters (1988) who has used econometric studies and 361 daily observations to analyse the connection between promotion and purchase behaviour of consumers. Similarly Dickson and Sawyer (1990) watched consumers in shops and their behaviour. They combined these observations with the results of a survey questionnaire to prove that buyers buy products advertised as special price items quicker than other products with regular price. Observation combined with questionnaire was used also in Murphy and Venkatesh (2006) when analysing the behaviour of prostitutes. Lovato et al. (2007) used observation to analyse the price kids are willing to pay for cigarettes. Cohen et al. (2007) observed alcohol purchasing behaviour. Ahmed (2008) studied time, atmosphere and activity between shopping. Nairn (2008) observed children aged 7 - 15 (video tapes) when using internet and analysed the advertising effects on children. Similarly Elliott (2009) analysed children when buying food in stores. Emmerton (2008) observed purchases in pharmacies and Krugman, Cameron and White (1995) observed consumers watching television in their homes.

By using the method of observation we aim to achieve measurement without cognitive bias and research bias. When developing a research model we argue that a relevant measure of stopping power should be executed in an as much as possible realistic environment. Only in a realistic environment we can measure the combined effect of attention (conscious) and engagement (subconscious) without having to control other research factors. So we define attention and engagement together as advertising stopping power.

Stating the advantages above, on the other hand, we acknowledge the limitations of this measurement. First the method of observation cannot measure the impact of multiple or 'mixed' emotions and is impractical for many marketing measurement purposes. Secondly (similar to eye-tracking) this measure is limited to the measurement of attention only. Finally if advertisers want to achieve overall effective advertising they need to achieve engagement or attention that can be seen in existence of stopping power. In the respect of the overall effectiveness of advertising we acknowledge the stopping power as a necessary but not self-sufficient measure. As it is only important when it leads to other positive effects, for example a change of consumer attitude and purchase intention, we discuss the content strategies in the following section.

Research Objective

The aim of the presented research is to demonstrate the usability of a method of observation for pre-testing. We plan to achieve this by analyzing

the differences in measuring stopping power for a 3D lenticular and standard printed ad. Thus we state the following hypothesis:

H1: stopping power measurement is a reliable measure (evaluators explain the same results in more than 90 % of cases and less than 10 % of similarities are due to coincidence).

H2 – the method of observation in attracting stopping power will provide valuable differences between two types of printed ads (printed in different techniques) for all locations although the content of the ad is the same.

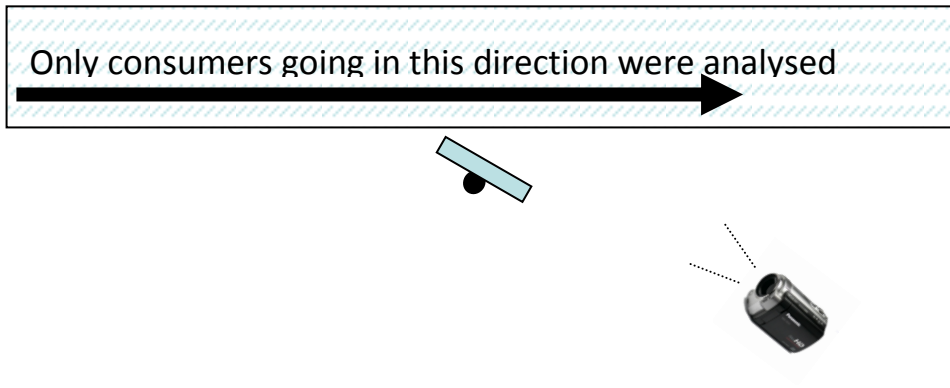
Research method

In order to test the method of observation we conducted a study to analyze how many people look at a poster in a shopping environment for more than 500 ms (half of a second). Direct exposure (a look at a poster at a 1-5 degrees visual angle) as suggested by Josephson (2005) was thus the only dependent variable in this research. In other words the poster had to stop the walking consumer. The interval 500 milliseconds should be long enough for inputs of adaptive importance to be perceived, to influence behaviour, and to be represented in the next stage of retention, short-term memory, while inputs without significance can disappear. Implications of this research are that consumers have enough time to recognize the brand name and a short message of practically any visible poster. Even though exposure time is short, the emotional and cognitive processes are unstoppable. Only after an ad has already been evaluated memory filter is applied. Thus although this measuring method of stopping power relies

solely on short-term effects it is the exposure itself that with repetition creates new patterns in consumer brain contributing to long-term brand recognition. Anchors to visual stimuli are namely unconscious.

The main idea of this research was to observe consumers in a real-time environment without any interaction with the researcher to make this evaluation as objective as possible. In addition we aimed to analyse only the differences in printing techniques to avoid the potential differences in the content of ads. Thus we created two posters of the same size (1.07 m x 0.85 m) and motive but printed in different techniques (standard 2D and lenticular 3D) and placed them according to Figure 1. We switched between the two posters every hour to satisfy the condition for comparison analysis and recorded behaviour using HD surveillance camera. Consumers did not know they were observed as the camera was a part of a surveillance camera used in shopping centre and marked only on entrance doors to the centre.

Figure 1: The setting of the poster and the recording camera.



Source: own research

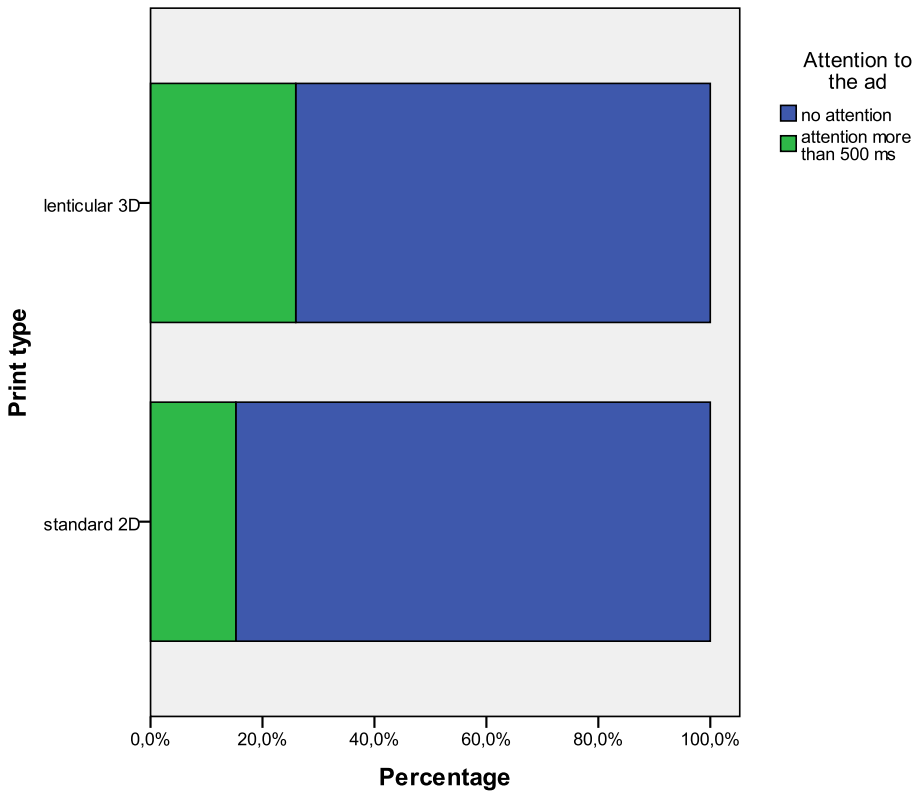
Results

Following the training for the two evaluators, who evaluated the research and were rewarded for this task in money, a total of 5115 consumer reactions to the poster were analyzed. Among these there were 2198 males (43 %) and 2917 females (57 %). For 2D ad 47 % and for 3D lenticular ad 53 % consumer reactions were analysed. Results from this study show that on average for 15.2% of consumers standard 2D advertisement attracted more than 500 ms of stopping power (visual attention). When using 3D lenticular poster the half a second limit of stopping power (visual attention) was reported with an outstanding 25.9% of consumers.

To test the hypothesis H1 tapes were analyzed twice in order to calculate the inter-coder reliability Krippendorff's alpha (α). The sampling

distribution of the means was assumed to be normally distributed as well as the sampling distribution of the scores. Based on the SPSS analysis described in Hayes (2007), Krippendorff's alpha showed fairly high reliability for all pairs of analyzed evaluations of the same consumers $\alpha = 0.935$. Alpha 0.935 means that 93.5 % of the units tested by evaluators are perfectly reliable while only 6.5 % are the results of chance. Reliability calculations using Krippendorff's alpha showed alphas at more than 0.90 level for all nine locations. Krippendorff (2006; 2011) suggest, for example, to rely on evaluators with variables $\alpha \geq 0.80$ although $\alpha \geq .667$ can already suffice for drawing tentative conclusions. Calculated alphas above 0.90 make this method a reliable variable for further analysis. Thus we can accept H1 that this method can be used to reliably measure the stopping power of consumers.

Figure 2.: Stopping power in advertising using 2D (standard print) and 3D (lenticular print)



Source: own research

In order to test hypothesis H2 we tested the reliability to measure differences between two ads on 9 different locations. Locations 1 to 7 were urban, locations 8 and 9 were in rural environment. All locations are in Slovenia. Percentages were similar for all 9 locations. Pearson's chi-square was calculated for differences between 2D and 3D for every location (table 1).

Table 1: Results stopping power on different locations

	% of Stopping power standard 2D ad	% of Stopping power standard 3D ad
Location 1 (Interspar, Ljubljana): $\chi^2 (1) = 4.438, p = .035$ *)	18.0	26.1
Location 2 (Interspar, Maribor): $\chi^2 (1) = 8.877, p = .003$ *)	15.7	26.4
Location 3 (Planet Tuš, Novo Mesto): $\chi^2 (1) = 10.990, p = .001$ *)	13.8	25.1
Location 4 (Hofer, Kranj): $\chi^2 (1) = 2.181, p = .140$)	15.9	20.4
Location 5 (Mercator, Nova Gorica): $\chi^2 (1) = 8.332, p = .004$ *)	14.8	23.6
Location 6 (Lidl, Koper): $\chi^2 (1) = 8.251, p = .004$ *)	16.0	25.0
Location 7 (Hofer, Domžale): $\chi^2 (1) = 29.475, p = .000$ *)	16.1	40.6
Location 8 (Mercator, Trzin): $\chi^2 (1) = 20.593, p = .000$ *)	14.3	30.6

Location 9 (Mercator, Ivančna Gorica): $\chi^2 (1) = 22.292, p = .000$ *	13.5	32.3
--	------	------

From the analysis we conclude that the ads tested attracted substantial stopping power. In our research stopping power was between 13.5 % to 18 % for the 2D ad and between 20.4 % and 40.6 % for the 3D ad. The difference between the two techniques is significant for 8 out of 9 locations. As the purpose of this study was to demonstrate the extent to which this method can identify the differences we can accept the H2.

Conclusion and further research

This research may serve as a starting point in the process of finding inexpensive, quick and efficient measurement tool in advertising. Additional research with similar methodology is however necessary to verify the results. Future research is especially recommended for new technologies that are seeking to find their way in advertising. With the introduction of 3D advertising (lenticular, holographic, 3D TV and other) the ability to stop the consumers will increase dramatically as we have seen when testing lenticular 3D where it almost doubled for the same motive of the poster (the difference was only the printing technique). In addition this methodology should be tested using different media.

Compared to traditional self-report measures this method is not influenced by cognitive bias. Moreover, due to no interference of researcher this

method provides a valuable realistic measurement tool that can be combined with other methods. Although it cannot measure complex emotions and fractionalises effects on different response categories it does include attention as well as engagement. It does not solve the problem of measuring long-term effects, however academic literature provides numerous links between exposure and long-term memory. This link is much stronger than traditional self-report measurement like recall or ad likeability. Finally this method gives instant feedback on how ads can stop the consumers and occupy their conscious or unconscious processing power.

Notes: This research was supported by the EU - Investing in your future – OPERATION PART FINANCED BY THE EUROPEAN UNION, EUROPEAN SOCIAL FUND.

References

- Ahmed, Allam (2008): Marketing of Halal Meat in the United Kingdom: Supermarkets versus local shops. *Food Journal*, Vol.: 110, No.: 7, pp.: 655–670.
- Berg, Bruce Lawrence (2001): *Qualitative Research Methods for the Social Science*. New Jersey, ZDA: A Pearson Education Company.
- Braun, Kathryn A., & Zaltman, Gerald (2006): Memory change: an Intimate Measure of Persuasion. *Journal of Advertising Research*, pp.: 57–73.
- Cohen, Deborah A, Schoeff, Diane, Farley, Thomas A, Bluthenthal, Ricky, Scribner, Richard, & et al. (2007): Reliability of a Store Observation Tool in Measuring Availability of Alcohol and Selected Foods. *Journal of Urban Health*, Vol.: 84, No.: 6, pp.: 807–813.
- Cradit, J. Dennis, Tashchian, Armen, & Hofacker, Charles F. (1994): Signal Detection Theory and Single Observation Designs: Methods and Indices for Advertising Recognition Testing. *Journal of Marketing Research*, Vol.: 31, No.: 1, pp.: 117–127.
- Del Vecchio, Eugene (1988): Generating Marketing Ideas When Formal Research Is Not Available. *The Journal of Consumer Marketing*, Vol.: 5, No.: 1, pp.: 65–68.
- Dickson, Peter R, & Sawyer, Alan G. (1990): The Price Knowledge and Search of Supermarket Shoppers. *Journal of Marketing*, Vol.: 54, No.: 3, pp.: 42–53.
- Elliott, Charlene D. (2009): Healthy Food Looks Serious: How Children Interpret Packaged Food Products. *Canadian Journal of Communication*, Vol.: 34, No.: 3, pp.: 359–380.

- Emmerton, Lynne (2008): Behavioural Aspects Surrounding mMedicine Purchases from Pharmacies in Australia. *Pharmacy Practice (journal)*, Vol.: 6, No.: 3, pp.: 158–164.
- Erevelles, Sunil (1998): The Role of Affect in Marketing. *Journal of Business Research*, Vol.: 42, pp.: 199–215.
- Hall, Bruce F. (2001): A New Approach to Measuring Advertising Effectiveness. North Carolina, ZDA: Howard, Merrell and Partners.
- Hall, Bruce F. (2004): On Measuring the Power of Communications. North Carolina, ZDA: Howard, Merrell and Partners.
- Hayes, Andrew F., & Krippendorff, Klaus (2007): Answering the Call for a Standard Reliability Measure for Coding Data. *Abingdon, VB: Communication Methods and Measures*, pp.: 77–89.
- Heath, Robert (2006): Emotional persuasion. World Advertising Research Center. Available at: <http://www.warc.com/> (12.5.2011), pp.: 46–48.
- Josephson, Sheree (2005): Eye Tracking Methodology and the Internet. V Ken Smith, Sandra Moriarty, Gretchen Barbatsis, & Keith Kenney: *Handbook Of Visual Communication Theory, Methods, And Media*. London, VB: Lawrence Erlbaum Associates., pp.: 63–81.
- Krippendorff, Klaus (2006): Testing the Reliability of Content Analysis Data: What is Involved and Why. Pennsylvania, ZDA: University of Pennsylvania. Available at: http://repository.upenn.edu/asc_papers/43 (12.5.2011)
- Krugman, Dean M., Cameron, Glen T., & White, Candace McKearney (1995): Visual Attention to Programming and Commercials: The Use of In-home Observations. *Journal of Advertising*, Vol.: 24, No.: 1, pp.: 1–12.

- Lovato, Chris Y, Hsu, Helen C. H., Sabiston, Catherine M., Hadd, Valerie, & Nykiforuk, Candace I. J. (2007): Tobacco Point-of-Purchase Marketing in School Neighbourhoods and School Smoking Prevalence: A Descriptive Study. *Canadian Journal of Public Health.*, Vol.: 98, No.: 4, pp.: 265–335.
- Nairn, Agnes (2008): "It does my Head in... buy it, buy it, buy it!" The commercialisation of UK children's Web Sites. *Young Consumers (journal).*, Vol.: 9, No.: 4, pp.: 239–253.
- Poels, Karolien, & Dewitte, Siegfried (2006): How to Capture the Heart? Reviewing 20 Years of Emotion Measurement in Advertising. *Journal of Advertising Research.*, pp.: 18–37.
- Shapiro, Stewart, MacInnis Deborah J., & Heckler, Suzan E. (1999): An Experimental Method for Studying unconscious Perception in a Marketing Context. *Psychology & Marketing.*, Vol.: 16, No.: 6, pp.: 459–477.
- Vakratsas, Demetrios, & Ambler, Tim (1999): How Advertising Works: What Do We Really Know?. *Journal of Marketing.*, Vol.: 63, No.: 1, pp.: 26–43.
- Young, Charles E. (2004): Capturing the Flow of Emotion in Television Commercials: A New Approach. *Journal of Advertising Research.*, Vol.: 44, pp.: 202–209.
- Young, Charles E. (2009): Ad Response Tests Show how Attention Connects to Memory. *Admap – World Advertising Research Center (WARC)*. Available at: <http://www.ameritest.net> (8.11. 2011), pp.: 42–45.

Zambardino, Adrian, & Goodfellow, John (2007): Being 'Affective' in Branding?. *Journal of Marketing Management*, Vol.: 23, No.: 1, pp.: 27–37.

Walters, Rockney G. (1988): Retail Promotions And Retail Store Performance: A Test Of Some Key Hypotheses. *Business And Economics–Marketing And Purchasing*, Vol.: 64, No.: 2, pp.: 153–180.