

Measurement of Advertising Effectiveness: How Different Theories About the Relative Importance of Sales vs. Recall/Recognition vs. Emotion Were Popular

Ana Beriain

Abat Oliba CEU University, Barcelona, Spain

The aim of this chapter is to run through the history of marketing research applied to measuring the effectiveness of advertising and deepen in all those theories that have shed light on this aspect and have influenced subsequent theories and models. This historical review offers the opportunity of knowing the past and understanding better the present of market research applied to advertising. It is a chronological analysis to present day, which introduces the development of the explanatory models for the functioning of advertising and the theoretical and empirical contributions in this area.

Keywords: advertising, models, measurement, effectiveness, market research

Introduction

This article provides an overview of the different stages that have historically taken place in market research applied to measuring the effectiveness of advertising. Through this historical review we see how the concept of effectiveness is developed and understood differently depending on the historical period or school of thought (European/American). We will also try to analyze in detail in each of the historical phases what variables may affect advertising effectiveness. Similarly, we will address the different tools and models that have been developed in the U.S. and Europe to measure objectively the effectiveness of advertising.

Objectives

Any advertising research should answer the question of whether the piece of advertising is effective in the sense that it meets the objectives and strategy for which it was created. To measure and evaluate the effect of advertising, the first thing to understand is how it affects individuals. Much effort has been devoted to this in the past and still today by psychologists, sociologists, scientists in different fields as well as marketing, advertising, and market research professionals. This analysis has led, over time, to different theoretical or explanatory models of how advertising works. Research institutes have incorporated and applied them across different types of studies, often radically different one from another. Therefore, and since in principle each system for measuring the effectiveness of advertising responds to a different way to understand how it affects individuals, it is necessary to examine these models in greater detail.

Ana Beriain, Ph.D., Associate Professor of Communication Department, Abat Oliba CEU University.

Correspondence concerning this article should be addressed to Ana Beriain, Universidad Abat Oliba CEU, Bellesguard, 20. 08022 Barcelona, Spain. E-mail: Beriain1@uao.es.

Stages by Means of Results

In the history of advertising several market research phases may be distinguished applied to the measurement of advertising effectiveness.

The First Phase: Sales as the Sole Objective of Advertising

A first phase highlighted by Hopkins contributions¹. He introduces research as part of the process of what he called scientific advertising; he was the first advertising man to make use of sampling and test markets. In 1923 he wrote *Scientific Advertising*, it is a treatise on advertising principles based on exact science and fundamental laws (Hopkins, 1945). Hopkins was a defender of research in advertising, and believes that most national advertising is left unaccounted for. It is simply assumed it is worth it, while a small test could indicate how to boost profits, and he stresses the importance of recollecting information before designing the advertising piece and after. Hopkins argues that the sole purpose of advertising is sales and that it is based on the same principles. Whether we agree or not with Hopkins, this led him to make a thorough analysis of the effects provoked by the direct advertising he created, and he did so by comparing the response of each action. He measured the direct response generated by each sales letter. This model applies the premise that the sole role of advertising is to sell and it is based on stimulus-response theories from the behaviorist school of psychology in the U.S.. The model examines advertising effectiveness through its effect on sales and, as we shall see further on, in the 1980s this led to the rise of research techniques based on econometric and statistical models aimed at measuring the increase on sales produced by each advertising action. The main criticisms of this measurement model, based on the direct relationship between advertising and sales, are that it does not take into account the mid and long term effects of advertising on sales nor the existence of intermediate variables that may influence the consumer between receiving the advertising stimulus and the purchase (Alloza, Both, & Benito, 2000).

The Second Phase: Recall, Recognition, and Persuasion, New Advertising Effectiveness Variables

A second phase of advertising evaluation emerges in the 1930s when methods start to be used based on advertising recall and recognition measures. In 1932, Starch, which look at the studies carried out by Gale about attention in advertising, measures the recognition of print ads. He uses two joint measures: having seen the ad and having read it. These two measures depend on the type of product, the market share of the product category and the communication media characteristics such as size of the ad, color, position, and focus of the copy. The validity of this system is questioned by Gallup since, according to him, the respondents can declare having read the ad only to please the interviewer or through confusion with a previous ad of the brand (Aaker & Day, 1991). Gallup, who never doubted the validity of scientific sampling or its potential value to society, developed a memory measure based on the premise that an ad was not effective only if recognized by the consumer, but should also be spontaneously recalled unaided.

Later in the 1950s begin the persuasion analyses methods as a truly indicative variable of effective advertising and in the 1960s Buzell presents the results of a series of multiple regressions using the Schwerin model (Buzell, 1964). Through this model, initiated by Malcolm and Murphy from The Schwerin Research Corporation and conducted by Murphy and Buzell, the psychological testing methods are applied to advertising analyzing advertising persuasion in television commercials. The pre-test based on the Schwerin model relates

¹ We cannot forget of the previous contributions of Scott in his book *The Psychology of Advertising* published in 1908 where he highlights the role of memory, attention, feelings, and emotions in advertising; and the contributions of Gale on the psychology of advertising (1900).

market shares to advertising expenditure and the quality of the message (content and presentation). He argues that the quality of the message is more important than expenditure in producing short-term changes in market share. The application of this model does not leave the advertising research industry indifferent and draws criticism as expressed by Fothergill and Ehrenberg (Fothergill & Ehrenberg, 1965).

At this time the Day After Recall (DAR) test also emerges, proposed by Procter and Gamble to the Burke research institute in Cincinnati. The methodology consists in phoning a sample of between 150 and 300 individuals who the previous day had seen the TV program during which the commercial to be analyzed had been aired. The main questions are based on the recall of the advertised brands and content (Aaker & Day, 1991). This type of test is not practically carried out nowadays as it has been replaced by other more efficient measurement tools. Zielske maintains a critical attitude towards this type of test, basically by the drawbacks for measuring emotional ads as content verbalization is simpler in more rational level ads (Zielske, 1982).

In the 1960s in the U.S. emerge models based on attitude changes, on persuasion, prior to purchase. These models were very successful in the 1960s and 1970s, based on that, before purchase, advertising must change consumer attitudes towards the brand. The mechanism raised by these models is very simple: at the same time that the advertising stimulus impacts the consumer, this communication has a persuasive effect that generates a more positive attitude towards the brand or product. This improvement in attitude will be reflected in an increase of sales. These models measure the change in attitude generated since that is the intermediate effect sought by advertising.

The three models developed and known as the hierarchy of effects, persuasion, attitude change or models step by step are characterized by presenting the existence of intermediate variables between the advertising stimulus and sales. The best known models are DAGMAR (Awareness → Comprehension → Conviction → Action); AIDA (Awareness → Interest → Desire → Action), and Steiner-Lavidge (Awareness → Knowledge → Liking → Preference → Conviction → Action). These models for measuring advertising effectiveness are based on the premise that advertising influences each of the steps in the purchase process for a product or service. Research, therefore, must be able to measure the influence of advertising in the process to purchase, in its different stages (Lavidge & Steiner, 1961).

The most notorious criticisms to these three models come from Brown, Bird, and Ehrenberg. On the one hand, Brown points out that advertising when seen by the consumer does not produce changes in basic attitudes toward brands but that, in real life, improvement of attitudes and increase in purchase intention are associated to contact with the product, as sales move before change of attitudes (Brown, 1991). For Brown, watching advertising at home is basically a passive experience in which defense mechanisms are also activated. In fact, Brown designed a model based on advertising memorization, close to the principles of the cognitive school, materialized in the measurement tools developed by the company Millward Brown. Depending on the model of advertising memorization, advertising introduces memories in the long-term that will affect later attitudes and consumer behavior towards the brand or product advertised. That is, advertising does not produce changes in attitude at the very moment when the consumer is exposed to it, but this effect will occur later. Moreover, Bird and Ehrenberg argue that the proportion of consumers who are spontaneously aware of a brand and the proportion of those who report its usage vary together from brand to brand but in a non-linear manner. Overall, the results show that brands with above-normal advertising expenditure over five years tend to have a higher awareness than usage share, and vice versa, for those brands spending below-normal levels (Bird & Ehrenberg, 1966). Years later, these authors, together with Channon, published an article where they maintained that the

consumer attitude towards a certain brand depends on how recently this brand had been used (Bird, Channon & Ehrenberg, 1970).

The Third Phase: Continuous Analysis of Advertising Campaigns

In the 1970s a third phase emerges. Advertisers and research institutes see the need to address the effectiveness of a campaign as a whole and not element by element. The main contribution in this respect is done, in 1976, by the Millward Brown institute. This company sets forth the study called Advanced Tracking Programme (ATP) consisting in continuous measurement of the effects of advertising related to advertising expenditure of the client's brand and its competition. In this type of studies, besides analyzing advertising recall and recognition, the analysis of brand image is incorporated (Brown, 2007). This type of study is still commercially operative today, and besides this company other institutes have incorporated the idea of continuous monitoring of advertising with their own measurement tools.

The discussion of recall versus recognition is one of the oldest in advertising research and most research institutions offer some measure of recall in their advertising test methodology. Another historical debate has been the analysis of persuasion versus recall. Persuasion is another variable usually applied by research institutes (Biel, 1993). In 1983 Gibson analyzes the effectiveness of research techniques based on persuasion versus recall and clearly advocates for the former, as he barely finds any association between recall and persuasion measures, while there is evidence of the relationship between persuasion measures with sales (Gibson, 1983). Currently, research companies are divided into these two schools and they use very similar measures to Gallup's recall pattern or Schwerin's persuasion pattern. In 1994, Du Plessis studies in depth the debate between recognition versus recall. The author reviews the experiments carried out so far on the analysis of recall and recognition, and performs three experiments with a large sample of ads. His conclusions are that the debate between recall and recognition makes no sense because, as liking and persuasion, they measure different aspects. Recognition is stronger and less subject to variations than recall because it measures the presence of an ad in the memory and recall does it through the brand. That is, memory has accesses to this data through the brand (Du Plessis, 1994).

At this point, we can speak of two schools: one focused on impact and recall. Some institutes presented commercials within a block of several advertisements and asked the sample what commercial they remembered seeing. The practical problem of this school is that a person's attention at a test is not at all the same as that of a person watching TV in a real situation; another school, based on persuasion, practiced laboratory tests where the sample was invited to attend a theater preview of a television program and in between a series of commercials were aired. On leaving the theater they were asked a declaration of purchase intentions, considering that if before seeing the commercials they had selected a brand and after watching them they had changed their mind for another one, the commercial was successful. A drawback of this is that persuasion does not work the same for all brands nor for all product categories.

The Fourth Phase: Recall Model vs. Persuasion Model

The fourth phase, also from the 1970s, is characterized by the incorporation of scanning, database of a single source, econometric modeling to evaluate advertising, modeling of advertising relations (in terms of pressure) and criteria of effects, including sales. Econometric analysis basically originated in the U.K. and the U.S.. The objective of this type of study is to analyze the effect advertising has on sales by isolation from the effect of other variables such as price, distribution, etc.. To do this, econometric models measure product sales

when the first advertising flight is aired and the residual effect that it continues producing some time later. This technique is perfectly valid to identify the immediate effect on sales. However, it does not contemplate the long-term effects of advertising. These econometric modeling techniques result in methodologies of measurement, especially developed by media agencies. In this regard, it is important to underline Jones contribution.

Jones (1995) stands out for his analysis of the effect of advertising on short-term persuasion. The author establishes a ratio, which can be calculated for each brand, called Short Term Advertising Strength (STAS), calculated as the ratio between the percentage of individuals who purchased the brand having seen its advertising and those who have purchased the brand without having seen the advertising, based on a very recent period such as the day before or the previous week. The STAS for each brand can be calculated using panel data as a sole source or by polling. The use of panel data is more expensive but more reliable, as the ratios obtained have a tendency to be lower than polling. As we noted at this stage Single Source Data analysis of advertising is also developed. This technique consists in selecting a fixed sample of households living in a geographically isolated location and who purchase in the same establishments, so that the consumers are subject to the same conditions such as price, distribution, promotions, and product categories available. After selecting the sample it is necessary to integrate several sources of information coming from the electronic cash register scanners installed at the point of sale, together with purchase information from the households and the television channels received in those same households. This technique is also known as Electronic Market Test and was developed in the U.S. by the company Marketing Scan under the name of Behavior Scan. Nielsen later introduced it in several countries, GFK in Germany and Scantrack in the UK. This type of test is very effective to measure the sales response for launches of new brands or existing brands that have never done advertising before, since in both cases it is possible to measure the response of short-term purchase in the households participating in the study and it provides a good measure of how effective the ads are when tested with this system. However, for known and established brands the results have been disappointing in general, and it is not very realistic to expect detectable results for an established brand in a short period of time (Alloza et al., 2000).

In this environment publications arise with reflections, suggestions, and criticisms of how advertising works on consumers. First the contribution done by Gallup (1974) must be highlighted. Gallup concludes that advertising based on the study of consumer needs helps to choose more wisely the products purchased and also will be more effective in providing relevant information.

Later, Vaughn (1980), Research Director of Foote, Cone and Belding, developed a new advertising strategic planning model based on fundamentals of the advertising process. For this he relied on the traditional theories of advertising, consumer behavior models and the latest developments in advertising research. The new model is based on the design of a product classification matrix relating the consumer's involvement (high or low, x and y axes) and whether they are rational products (requiring detailed information) or emotional, abscissa axis. Thus, four quadrants are formed for the development of advertising strategies: informative strategy (rational-high involvement), affective strategy (emotional-high involvement), habitual strategy (rational-low involvement), and satisfaction strategy (emotional-low involvement). The final objective of any strategy is to be able to identify what we want consumers to feel, know, and do with respect to our product or service, and the model proposed by Vaughn helps standardize and rationalize the infinite possibilities in an organized and practical format. In 1986, Vaughn re-edits a new version of the matrix model for advertising strategic planning where he details in-depth the products and services that may appear in each quadrant, as well

as the differences between the four types of advertising strategies derived from the model. The model proposed by Vaughn harmonizes and puts together all these different theories with the objective of their culmination (Vaughn, 1986). It is at this point that the model becomes known as “The FCB Grid”.

As we have seen, the two most common models for advertising analysis, the recall model versus the persuasion model, that is the British model versus the American model, generated over the years various opinions among advertisers and research institutes. This was one of the reasons why Dun at the annual conference of the Advertising Research Foundation (ARF) in 1977 proposes to address the validity of studies to measure the effectiveness of advertising in a joint manner. Dun argues that each researcher had partial data, but by crossing the information the industry could learn from the experience. He called forming a committee to validate copy-testing of ads analyzed for clients with absolute confidentiality guarantee, which was called The ARF Copy Research Validity Project. However, this attempt was unsuccessful because after three years dedicated to the compilation of cases, those available were so few that did not lead to any meaningful conclusion, so another different method was devised that did prove successful. The results were presented at the ARF annual conference in April 1990 and were published later (Haley & Baldinger, 1991).

In 1990, the ARF Copy Research Validity Project establishes a Decalogue of everything that should be included in an optimum advertising analysis and concludes that both models are good, meaning that neither is entirely correct and that each model covers a different area of advertising.

Undoubtedly, the main finding of the study is the strong correlation between likeability of an advertisement and results in sales of that brand. However, this document has generated much literature on the subject: among which must be emphasized the critical views by Rossiter and Eagleson (1994) conducted on this study. The authors argue that many of the methodologies studied with percentages of correct sales prediction of 90% do not include likeability measures. Moreover, the authors believe that in their study it is demonstrated that persuasion is more crucial in forecasting sales. Brown also believes that likeability is one of the most important factors in predicting the sales results of an advertising campaign. He carries out a study showing that likeability exceeds persuasion and notoriety as a measure of prediction (Brown, 1991).

From this moment two relevant issues can be pointed out: the monograph published by Heath (2001) that will form part of a fifth phase, and finally the application of neuroscience knowledge to marketing, and in particular to advertising research. Neuromarketing begins to emerge as the new trend in advertising and market research.

The Fifth Phase: Analysis of Emotions

Heath calls into question the traditional models used in advertising research and shows how most advertising is processed at an emotional level (not a rational level). His investigations lead to a further review of market research approaches that compare the behavior of those who have seen advertising compared to those who have not. Heath argues that much of advertising is not articulated on recall, while Brown believes that recall is a major step in achieving efficiency. From this point, techniques are developed analyzing emotions provoked by the brands on the consumer. Some market research institutes already incorporate them into their methodologies for measuring the effects of advertising on consumers.

Heath (2009) argued that advertising in television can build big brands with low levels of attention through emotional engagement, a new definition of engagement independent of the concept of attention. He defines emotional engagement as the amount of subconscious feeling (emotional thoughts) going on when an

advertisement is being processed. The author dismantles the theory that engagement is related to attention. Heath defines attention as the conscious thinking flowing through while processing an ad while engagement is defined as the amount of subconscious emotional feeling going on at the same time. What is important of these two definitions is that they do not overlap. In other words, advertising can generate high attention but low engagement, and vice versa, high engagement with low attention. Quantitative research conducted by Heath suggests that advertising in television is particularly deficient in capturing attention, but especially efficient for engaging consumers. Not surprisingly, for Heath, television is the most effective media for building big brands.

The Sixth Phase: Neuroscience Applications

Finally, neuromarketing, which consists in the application of neuroscience to the field of marketing, tries to study the effects of advertising in the human brain. Neuroscience and how it applies in the recollection of information regarding emotions is a new trend in which most research institutes are working today. Neuromarketing rises with the boom of neuroscience in the 1990s and with the emergence and fine tuning of image diagnosis technologies such as magnetic resonance imaging (MRI) or magnetic resonance tomography (MRT), which allows to observe in images the structure and composition of the brain, and functional magnetic resonance imaging (fMRI), which allows to view in images the brain areas that become activated when performing certain actions. These imaging techniques are combined with other technologies that allow analysis of the effects that cause many visual and auditory stimuli. Through these techniques, scientists can perceive neural activity by measuring the brain's blood flow and electric impulses in each of its areas during information and advertising processes. However, these techniques are still in the making and many professionals are wondering if these modern techniques are the definitive answer to measuring advertising effectiveness.

In the area of advertising there can be found both defenders that the rational level of an ad is the best vehicle to sell products better, as well as those who argue that emotional level advertising reduces the reluctance of the client with respect to price and enhances brand differentiation. In this sense, greatest knowledge of the consumer's emotions will be the best for more effective advertising.

As can be deduced on reading this article, there is no consensus as to which model is best suited for measuring the effect of advertising. However, where the different agents involved in the advertising process do seem to agree (advertising professionals, researchers, advertisers) as well as other types of professionals (psychologists, sociologists, anthropologists) is in the different levels of response from the purchaser: cognitive, affective, and conative.

Conclusions

In view of the different approaches exposed regarding the effectiveness of advertising in the historical periods analyzed, it can be concluded that:

Throughout the history of advertising, efforts have been dedicated from the professional and academic talents to develop a model capable of measuring advertising effectiveness in order to analyze as objectively as possible the return over advertising expenditure.

Effective advertising is a multifaceted and multivariable concept (linked to sales, persuasiveness, brand awareness, message comprehension, recall, recognition, type of argument (emotional/rational), likeability, attention, involvement, credibility, and coherence), which conveys its approach from an interdisciplinary point of view (psychology, communication, anthropology, neuroscience, etc.). This complexity obliges to evaluate

advertising effectiveness in relative rather than absolute terms.

The emergence of new media and advertising formats with the existence of new communication channels added to the acquisition of internet skills by consumers (increasingly active and selective in the exposure to advertising) requires continuous reformulation of models and tools for measuring advertising effectiveness.

It can be affirmed that the systems for measuring advertising effectiveness developed by market research companies nowadays are capable of quantifying diverse specific aspects of advertising. With the objectives aimed in advertising it is important to establish, from the beginning, what variables are to be measured in order to select the most adequate tool or measurement system.

References

- Aaker, D., & Day, G. S. (1991). *Market research (Investigación de mercados)*. México: McGraw-Hill.
- Alloza, A., Both, C., & Benito, M. (2000). General framework: How TV advertising works? (Marco general: ¿cómo funciona la publicidad en televisión?). In J. Martínez et al. (Eds.), *Research in marketing (La investigación en marketing)* (pp. 819-857). Barcelona: AEDEMO.
- Biel, A. (1993). *Ad research in the US.: Hurdle or help? A brief review of the state of the art*. Warc: Admap Magazine.
- Bird, M., & Ehrenberg, A. S. C. (1966). Non-Awareness and Non-Usage. *Journal of Advertising Research*, 6(4), 4-8.
- Bird, M., Channon, C., & Ehrenberg, A. S. C. (1970). Brand image and brand usage. *Journal of Marketing Research*, 7(3), 307-314.
- Brown, G. (1991). *How advertising affects the sales of packaged goods brands: A working hypothesis for the 1990's*. Warwick: Millward Brown International Plc..
- Brown, G. (2007). *How tracking studies can help measure effectiveness*. Retrieved January 10, 2007, from <http://mbonline/kbank/referencepapers/nativepapers/trackingstudies.doc>
- Buzell, R. D. (1964). Predicting short-term changes in market share as a function of advertising strategy. *Journal of Marketing Research*, 1(3), 27-32.
- Du Plessis, E. (1994). Recognition versus recall. *Journal of Advertising Research*, 34(3), 75-91.
- Fothergill, J. E., & Ehrenberg, A. S. C. (1965). On the Schwerin analysis of advertising effectiveness. *Journal of Marketing Research*, 2(3), 298.
- Gale, H. (1900). Gale, H. (1900). On the psychology of advertising. In H. Gale (Ed.), *Psychological Studies*. Minneapolis, MN: H. S. Gale.
- Gallup, G. (1974). How advertising works. *Journal of Advertising Research*, 14(3), 47.
- Gibson, L. D. (1983). Not recall. *Journal of Marketing Research*, 23(1), 39-46.
- Haley, R. I., & Baldinger, A. L. (1991). The ARF copy research validity project. *Journal of Advertising Research*, 31(2), 11-32.
- Haley, R. I., & Baldinger, A. L. (2000). The ARF copy research validity project. *Journal of Advertising Research*, 40(6), 114-135.
- Heath, R. (2001). *The hidden power of advertising*. London: Admap Monographs.
- Heath, R. (2009). Emotional engagement: How television builds big brands at low attention. *Journal of Advertising Research* 49(1), 62-73.
- Hopkins, C. (1945). *My live in advertising (Mi vida en la publicidad)*. Barcelona: Aleu & Domingo.
- Jones, J. P. (1995). *When ads work: New proof that advertising triggers sales*. New York: Lexington Books.
- Lavidge, R., & Steiner, G. (1961). A model for predictive measurements of advertising effectiveness. *Journal of marketing*, 25(6), 59-62.
- Rossiter, J. R., & Eagleson, G. (1964). Conclusions from the ARF's copy research validity project. *Journal of Advertising Research*, 34(3), 33-34.
- Scott, W. D. (1908). *The psychology of advertising*. Boston: Small, Maynard.
- Vaughn, R. (1980). How advertising works: A planning model. *Journal of Advertising Research*, 20(5), 27-33.
- Vaughn, R. (1986). How advertising works: A planning model revisite. *Journal of Advertising Research*, 26(1), 57-66.
- Zielske, H. A. (1982). Does day-after recall penalize "Feeling" ads? *Journal of Advertising Research*, 22(1), 19-22.