

Chapter 23

AD TESTING

Allan L. Baldinger, Solomon & Associates

William A. Cook, Advertising Research Foundation

Historical Context

Growing in Use and in Controversy

Advertising research grew dramatically as the budgets for advertising with the birth of television advertising, but declined as a share of research spending in the 1980's as scanner data turned the attention of marketers to the short-term gains driven by trade promotions. Whether high times or low, Ad Testing has been frequently embroiled in controversy.

David Ogilvy and Leo Burnett were great ad men and great advocates of advertising testing. In 1994, David Ogilvy puzzled, "Most creative people today detest research, and I've never understood why.... In my day, I used research very often to give me the courage to run campaigns that were risky. My famous eye-patch ads for Hathaway shirts, for example, would not have been created had I not been studying a chart that I saw in Harold Rudolph's *Attention and Interest Factors in Advertising*."

Today many creatives would still differ with Ogilvy and take the position of Kevin Roberts, CEO, Saatchi & Saatchi Worldwide, who counters, "...two of the biggest mistakes we make in advertising research involve copy testing and the timing of the research. Despite what our creatives will tell you, not all research kills creativity, but quantitative copy testing certainly does. It does so by attempting to quantify the unquantifiable. Roberts urges creatives to avoid the

“overuse of research as a form of judgment on creative work, and its under use as a source of insight into the mind and mood of the consumer.”

The creatives’ hostility towards advertising testing grew particularly heated in the late 1970’s when marketers pushed for simplification and a single-measure on which to make copy-selection decisions. The “one-number-fits-all” mentality was a sharp stick in the eye of creatives seeking to tailor their ads to particular groups of consumers and competitive situations.

Charles Young (2004) has observed that “there are four general themes woven into the last half-century of copytesting. The first is the quest for a valid single-number statistic to capture the overall performance of the advertising creative. These are the various “report card” measures used to filter commercial executions and help management make the go/no go decision about which ads to air. The second theme is the development of diagnostic copytesting, whose main purpose is optimization, providing insights about and understanding of a commercial’s performance on the report card measures with the hope of identifying creative opportunities to save and improve executions. The third theme is the development of non-verbal measures in response to the belief of many advertising professionals that much of a commercial’s effects — e.g., the emotional impact — may be difficult for respondents to put into words or scale on verbal rating statements and may, in fact, be operating below the level of consciousness. The fourth theme, which is a variation on the previous two, is the development of moment-by-moment measures to describe the internal dynamic structure of the viewer’s experience of the commercial, as a diagnostic counterpoint to the various gestalt measures of commercial performance or predicted impact.

The Ad Testing being conducted today has evolved and grown substantially since World War II. There is a direct correlation between advertising expense and the expenditures on Ad

Testing. Since the bulk of advertising expense is now devoted to Television, Ad Testing research in the U.S. is predominantly conducted on Television ads.

Growth of the Media

Advertising Testing followed the development of the Media. As Media spending changed over time, so did the research associated with it. Prior to World War II, the predominant Media in the U.S., was Newspapers, Magazines (or Print), and Radio. But, when Television exploded on the scene in the 50's, advertising research and testing exploded along with it.

Table 23.1 lists some publicly-available figures on the relative importance of the various media, in the U.S., for the first half of 2004. The total spending was almost \$68 billion, which would mean that, on an annual basis, Ad spending in the U.S. is roughly \$140 billion.

Table 23.1

Media Spending

<u>Media</u>	<u>\$ Spending (\$ bill.)</u>	<u>% of Total</u>
Television	\$29.6	44%
Newspapers	\$14.4	21%
Magazines	\$13.2	20%
Radio	\$5.3	8%
Internet	\$3.7	5%
Outdoor/billboards	<u>\$1.4</u>	<u>2%</u>
TOTALS	\$67.6	100%

(Source: TNS Media Intelligence/CMR)

The fastest-growing of these media is Internet advertising, which grew at a + 26% rate vs. the first half of 2003. Still, advertising on the Internet represents only 5% of total ad spending. There are no publicly available figures on how the Ad Testing business itself splits by media. However, it is likely that the proportions differ somewhat from the spending figures shown here. For example, most Newspaper advertising is really retailer-based sales or promotion-related ads. These are rarely tested with consumers by the advertisers who buy them, since they are merely intended to spur a short-term sales boost, not to communicate an enduring or persuasive sales message. Many newspaper ads are sponsored by small local companies, which lack the financial resources to conduct sophisticated Ad Testing.

Additionally, many ad campaigns start out as TV campaigns, where the bulk of campaign money is spent, then are spun off into other media as Magazines or Radio ads using the same themes and content as were used in the TV campaign. Consequently, many advertisers test only the TV campaign for its likely effectiveness, and trust that its effectiveness will be easily translatable into other media. This assumption of equivalent effectiveness in other media as in TV has been found wrong in numerous instances.

Still, many campaigns are specifically oriented towards, and tested within, the non-television media outlets, such as Radio, Print or the Internet, as will be discussed below.

Evolving Research Philosophy

Although a change appears underway today, Ad Testing remains dominated by a model called the AIDA model, originally published as an explanation of how personal selling works by St. Elmo Lewis, in 1898. The AIDA model stands for:

A Attention

- I Interest
- D Desire
- A Action

The AIDA model is part of the Hierarchy of Effects theory, which holds that the basic and ultimate objective of all advertising is Sales. It is a linear theory, meaning that it presumes that consumers must go through a rational and sequential series of steps. The task of advertising begins with developing “attention”, or awareness of the brand being advertised. The second stage is to generate “interest”, through the communication of a relevant sales message. The third stage is to create “desire”, the ability to persuade the consumer that the communication is convincing enough that she is actually motivated to buy the brand. And the fourth stage is “action”, usually the actual purchase of the brand being advertised.

It should be noted that sales is not always the ultimate goal of advertising. For example, in automotive advertising, the goal is to create sufficient interest to motivate consumers to visit a showroom. The actual sale takes place once the retail salesman takes over, and following a test drive. Similarly, Public Service Advertising often has as its goal a simple shift in consumer attitudes (for example, Smokey the Bear motivated consumers to “prevent forest fires” by being careful with matches in National Parks, while anti-smoking campaigns are intended to reduce the sale of cigarettes, not increase them). However, in virtually all cases, advertising has its goal at least a shift in “Desire.”

While the outlines of the AIDA hierarchy are widely visible in most approaches to ad testing today, the hierarchy has grown in layers, variations and complexity. As early as 1974, Michael Ray (1974) noted that there are three types of behavior exhibited in consumer behavior: cognition, affection and conation, or thinking, feeling and doing. Batra and Ray (1986) found that in different

situations, the order of occurrence of those behaviors varied – that feelings could precede, and therefore influence, thinking and doing.

The great volume of brain research over the last decade has pushed the interest in emotion and affective response to the front stage. Gerald Zaltman (2003) went so far as to declare that senior managers need to change the way they think about how their customers think:

“The most troubling consequence of the existing paradigm has been the artificial disconnection of mind, body, brain and society... Only by reconnecting the splintered pieces of their thinking about consumers can companies truly grasp and meet consumers’ needs more effectively—and thus survive in today’s competitive and rapidly shifting business environment.”

Eight Ad Testing companies are participating along with Zaltman and leading advertisers and agencies in an ARF-AAAA consortium study of the emotional response to advertising. One of the participants is John Hallward from Ipsos-ASI, a long-time provider and avid defender of recall testing. Hallward (2005) observes, “Since the first impressions of advertising may often be more emotional than rational, we need to explore beyond the rational to better understand consumers’ emotions towards and impressions of the product or the message.

Models

Models of How Advertising Works

The AIDA model described above has been integrated into numerous models in current use today. They are broadly referred to as Hierarchy of Effects models. As the constructs of consumer attitudes and emotions were added, the models have become more than frameworks for teaching about advertising and have gained increasing relevance to measurement of advertising.

Bill Wells (1989) observed that ad forms could be described as “lectures” or “dramas” or combinations of the two. A lecture is an ad in which an announcer speaks directly to the viewer through the TV screen. These ads tend to do well on persuasion, but less well on attentiveness. Most demonstrations of a brand’s specific functional benefits are lectures. On the other hand, an ad which is categorized as a drama is a mini-play. An ad of this type can be highly entertaining, and can do very well on measures of attentiveness. However, the creative challenge in these ads is often to ensure that relevant brand benefits are also properly communicated. Wells also introduced an important distinction between the motivational routes of the advertising; rather than simply being informational, ads could be transformational, for example, enhancing the product consumption experience or the esteem one feels when wearing or using the brand (Puto & Wells, 1984).

For the ad agency Foote Cone Belding, Richard Vaughn (1980) developed an attitudinal model, the FCB Grid, it tied together the two constructs, High Involvement/ Low Involvement and Thinking/Feeling, into a simple 2X2 grid. Products such as insurance and televisions were placed in the thinking/high-involvement quadrant, while salty snacks and greeting cards went into the feeling/low-involvement quadrant.

Rossiter and Percy (Rossiter, Percy, & Donovan, 1991) expanded on the FCB grid. They added brand awareness in their model and they replaced the Thinking/Feeling distinction with the type of motivation involved in the ad – either informational or transformational.

For transformational motives they included sensory gratification, intellectual stimulation and social approval.

Media Models

In 2002, the Advertising Research Foundation published a monograph entitled “Making Better Media Decisions” updating the 1961 version, “Toward Better Media Comparisons.” These papers described the various stages in which media is communicated and measured, described below. The purpose of the 2002 update was to reflect the changes in the media environment which had occurred between 1961 and 2002, such as the introduction of the Internet as a new media vehicle, and the increased emphasis on Sales.

Both models include, as key initial stages, Vehicle Distribution, Vehicle Exposure and Advertising Exposure. Both models also include Sales Response as the ultimate and final purpose of advertising. However, the updated model changes Advertising Perception to Advertising Attentiveness, and adds Advertising Persuasion and Advertising Response. In the 2003 model, the first 3 are descriptions of the media and its distribution, while the last 5 describe consumers’ reactions to the media.

Here are the definitions of each of the 8 stages, taken directly from the 2002 monograph (available from the ARF at www.theARF.org)

1. **Vehicle Distribution.** This is a count of physical units through which advertising is distributed. It is a pure media effect. Measurement techniques include newspaper and magazine circulation studies, TV and radio tuning studies, online-media page requests, billboard locations.
2. **Vehicle Exposure.** This is a count of the people exposed to the media vehicle whose eyes or ears are open. It too is a pure media effect. Measurement techniques include radio and TV-people ratings, magazine-readership studies, online media page-view counts, billboard traffic counts, etc.

3. **Advertising Exposure.** This is a count of the people exposed to the media vehicle who are also exposed to its advertising. It is the highest level of measurement that is still a mostly pure media effect. Measurement techniques include radio and TV commercial-audience ratings, print ad page-exposure studies, online ad-view counts, billboard-traffic counts, etc.
4. **Advertising Attentiveness.** This is the degree to which those exposed to the advertising are focused on it. It is the first measurement level at which the effects of the medium are significantly confounded with the effects of the creative. Measurements include dedicated attentiveness studies, recall or campaign tracking studies, brainwave research, etc.
5. **Advertising Communication.** This is a measure of the information retained by the consumer after exposure to the message. Measurement techniques include advertising and brand awareness tracking, copy testing/recall, advertising recall studies, etc.
6. **Advertising Persuasion.** This is a measure of the shift in intentions produced by Advertising Communication. Here we are interested in the medium's ability to frame the message in ways that make it more credible, more relevant, and hence more persuasive. Measurement techniques include advertising tracking, copy testing, intent to purchase, willingness to consider, etc.
7. **Advertising Response.** This refers to measures of consumer response short of sales. Examples include visiting a showroom, calling a toll-free number, clicking on an online ad, requesting a brochure, etc. In direct mail and interactive media, such responses can be measured directly. Measurement techniques include click-through, post-click-through interaction, lead generation, telephone and mail response, coupon redemption, etc.
8. **Sales Response.** This is purchase of the advertised product or service in response to the advertising. Of all the measures listed, it is the most relevant to the advertiser, but the least

dependent on advertising and media effects. In addition to sales, useful measures include profits, Return on Investment (ROI) and Consumer Lifetime Value (CLV) – an estimate of the future profitability of a newly-acquired customer. Measurement techniques include sales tracking, test markets, single-source panel research, and marketing mix modeling.

The ARF Media Model provides advertisers, the media, ad agencies, and research companies a useful framework to address the research complexities and issues involved in the measurement of advertising effectiveness.

Table 23.2 provides the types of research most relevant to the individual stages. Stages 1 through 3 are addressed by media research measures of media delivery and audience exposure. Stages 4 through 6 are the areas where Copy Testing techniques are applied, while stages 7 and 8 are those for which Ad Tracking and In-Market Sales Tracking methods are used.

The industry has long desired the availability of techniques that provide a “single source” method of advertising measurement, e.g., a single system that links media or advertising exposure to sales response.

Table 23.2

Types of Research

2002 ARF Media Model	Research Type
1. Vehicle Distribution	Media Research
2. Vehicle Exposure	Media Research
3. Advertising Exposure	Media Research
4. Advertising Attentiveness	Ad/ Copy Testing

5. Advertising Communication	Ad/ Copy Testing
6. Advertising Persuasion	Ad/ Copy Testing
7. Advertising Response	Ad/ Copy Testing/ In-Market Tracking
8. Sales Response	In-Market Tracking

Advertising Effects Models

In the context of marketing research, a “model” typically refers to an analytic approach to a given dataset, designed for the purpose of adding value to the data. Models vary widely in complexity, approach and scope. They range from purely descriptive models, such as the ARF Media Models just described, to highly technical and sophisticated models employing advanced statistical techniques.

Sometimes the same research project or study will provide useful measures of several stages of the ARF Media Model. However, measuring the effectiveness of a particular advertising campaign is not a simple task. Since a given brand may be spending money on multiple media outlets and advertising campaigns simultaneously, as well as event sponsorships, PR, trade and customer promotion, teasing out the actual ROI of an advertising campaign or expense can be very difficult.

In order to deal with this complexity, many of the research companies involved in the measurement of advertising have developed their own models of advertising effectiveness. These can vary considerably in their complexity, and from company to company. Indeed, sometimes the same research company will employ different modeling approaches depending on whether a given

research project is a copy pre-test, a survey-based ad tracking project, or an in-market evaluation of long-term spending vs. sales trends.

There is considerable published literature covering the validation efforts by various research companies, much of which discusses the general modeling approaches used (e.g., Adams & Blair, 1992; Lukeman, 1995). However, the technical details of the various models often remain within the research company, as a proprietary asset of the company. These technologies are frequently a key to the company's differentiation in the crowded Ad Testing marketplace.

The Stages of Advertising Research

Developmental Research and the Qualitative Role

Research is performed at the various stages of the advertising process. However, it is estimated that well less than half of the ads in the media each day have been developed through a rigorous process of research..

Research can contribute to advertising at several points of development:

- Copy Development stage
- Rough Commercial stage
- Final Production stage, prior to air
- After Airing

The **Copy Development** stage is the initial stage in the development of new advertising of a product or service. Often, Ad Agencies talk about this stage as a Copy Exploratory. A very common research method used at this stage is focus groups, or other forms of qualitative research. Small

groups of targeted consumers are recruited, usually by phone, and are brought into central location facilities. This is the stage at which the client's ad agency is most-directly involved. In fact, many copy exploratory research projects are paid for by the ad agency.

The main purpose of copy development/ exploratory research is to help identify the optimal advertising strategy to employ in the new advertising. The strategy represents the underlying communication objective of the ad. It represents a summary of the claim or claims to be communicated about the brand, and the supporting evidence for that claim. The execution is the advertising agency's creative interpretation of how to most-effectively translate the selected strategy into an actual advertisement.

A typical Strategy Statement for an ad campaign might look something like this:

“The (BRAND) is better at (CLAIM/BENEFIT), relative to (COMPETITIVE FRAME), because it (SUPPORTING EVIDENCE).”

A completely hypothetical example might be:

“Rolaids is better at curing heartburn within 20 minutes, relative to Tums, because it contains a new and improved active ingredient.”

In a typical developmental qualitative project, a variety of possible claims, and/or supporting evidence, will be shown to possible brand buyers, and/or current brand buyers. Insights will be obtained about the relative merits of these benefit statements, product claims, and potential supporting evidence. It is usually not advisable to settle in on one final strategy or claim, based on group sessions alone. That is because group sessions, by their very nature, employ small sample sizes of consumers, and because a variety of possible biases can creep into such sessions. One type of bias of concern to those using focus group research is *leadership bias*, which occurs when the

session's respondents merely agree with a particular strong or articulate group participant. Another is *social acceptability* bias, in which case, respondents voice opinions based on social norms, rather than their true feelings (e.g., "I never buy any brands based on advertising.").

The general purpose of such qualitative sessions is to reduce the number of possible strategic alternatives, to a smaller more-easily-testable number of alternatives.

Many research companies have also created quantitative *early stage* approaches. In these techniques, alternative advertising strategic approaches can be ordered on their predicted effectiveness based on large sample sizes of consumers, who are exposed to advertising stimuli in the form of simple statements of claims or as rough print ad style concepts.

Evaluative Pre-Testing

Once the copy exploratory developmental phase has been completed, the marketer will usually move in the direction of testing the advertising, using quantitative methods. This is often called *copy pre-testing*, with the "pre" referring to the testing of advertising **prior to** actually placing the advertising in the media.

Depending on whether or not the number of advertising executions is large or small, and whether the likely costs of production of the advertising is high or low, the client and ad agency may either test the advertising in "rough" or "finished" form. In other words, an advertisement may be inexpensively simulated by means of a "rough" version of the advertisement.

The Value of Norms

Once an advertiser decides which ad-testing measures are likely to be the most predictive of marketplace success for his brands, normative databases become extremely valuable. She knows

that some of his ads have been more successful than others in the past, and often believes that ad campaigns must be frequently changed because of "wearout". But she also knows that advertising success will vary by product category, and even from brand to brand within the same category. Therefore, a database of scores, gathered consistently over time, using the same technique and sample specifications, is the best means of separating better ads from weaker ones (Walker, 2002).

Such norms can become key strategic assets for copy-testing companies, since they are the source of much of their R&D, and can form the basis for their claims of predictive validity. Category-specific norms are regarded by client advertisers as being more valuable than the overall normative databases of the company. However, the norms which are most predictive are those which are specific at a brand level. Because a norms-based system can lead to years of loyal use of a given copy-testing technique by a client, norms can also be central to the basic business model of ad testing companies.

Ad Testing Formats

There are a variety of formats of rough advertisements used in testing. Some of the more-predominant forms are:

- Animatics
- Photoboard/storyboard-based simulations
- Live action roughs

Animatics use cartoons to simulate the visuals of the advertising. Photoboards use early or stock-based photos to recreate the intended visuals. Live-action roughs use live actors and sets, but attempt to compromise on the expenditures that would be used to produce the final commercial or commercials.

Usually, when advertising is quantitatively tested in rough form, the intent is to screen a larger number of commercial alternatives down to a smaller number of commercials. Once this screening test is conducted, the winning commercial or commercials will typically be produced in final production, and then re-tested. All of these tests, at both stages, qualify as ad pre-tests, inasmuch as all such tests are still conducted prior to going on air.

Objectives of Pre-Tests

The basic purposes of ad pre-tests are three-fold:

- to select the optimal commercial for actual media placement
- to determine whether or not there are specific components of the ads that need to be altered or improved, and
- to determine whether or not the optimal advertisement is likely to perform in a superior manner to advertising that is currently on air. (NOTE: This last objective is obviously not relevant in the case of pre-tests being conducted for new products, which have not as yet received any live advertising exposure.)

However, because of severe time pressures in today's ad-testing environment, ads are often tested after they have already begun to run on air, or are not tested at all. An unknown question is the proportion of advertising campaigns which are receiving air time/ significant media, but have not been exposed to consumers for test purposes prior to running on air. If this trend were to continue, the advertising research industry will receive increasing levels of justifiable criticism for a lack of standards of accountability.

Post On-Air Ad Tracking

Once the final advertising execution has been tested, and possibly refined, and retested, it is placed in the media. Decisions are simultaneously being made by the client with respect to how much money will be spent on the ad campaign, whether the campaign should consist of one execution alone or multiple versions or *poolouts*, whether the poolouts themselves should also be tested, how the advertising campaign budget should be allocated across the media alternatives, how the budget should be spent across time periods, and whether the versions of the advertising campaign to be placed in alternate media vehicles should also be tested.

For example, it might be the intent of a given campaign to be aired with a certain proportion of the budget allocated towards a TV campaign, say 60%, but that the other 40% be spent on a mix of Print Ads, Radio Ads, and/or Online banner ads. Each of these media alternatives require slightly different Ad Testing techniques, to be discussed below.

Once these decisions are made, and the campaign is aired, another research project may be undertaken, called Post On-Air Ad tracking.

Pre tests are individual research projects conducted at a given point in time. However, Ad Tracking projects attempt to measure the effectiveness of the Ad campaign over the course of time. A typical ad tracking project is an annual contract, calling for interviews to be conducted over the course of a year. In the case of an ad tracker for a new product, or if an ad tracking project has not already been set up to monitor the advertising in the client's category, the project may include both "pre" and "post" interviews, conducted both before and after the ad begins to run on air.

Projects may be conducted using a variety of data collection methods, such as telephone, online, mall intercept, or mail interviews. Over the past 20 years or so, there has been a shift between periodic or wave-based interviews for such projects, towards continuous interviewing. In a wave-based project, a set of interviews might be conducted prior to the onset of advertising, perhaps

consisting of 300 to 800 category users, conducted the month prior to the beginning of the ad campaign. Additional interviews would then be conducted at periodic intervals following the campaign's onset, perhaps at 3 months following the campaign's onset, 6 months, 9 months, and 12 months following the campaign's airing. The design of the project would depend on a variety of factors including the expected media weight, and the extent to which it is evenly spread over the months of the campaign, or will be front-loaded towards the early months, or backloaded towards the end of the year. The seasonality of the category and brand is another factor affecting the research design.

The potential problems with wave-based ad tracking projects are that it is impossible to predict how or when competitors may be simultaneously changing their media spending, or strategies, and that the actual peaks or valleys in the campaign's effectiveness could easily occur during the periods in which the interviews are not being conducted.

Continuous ad trackers include an interviewing design that calls for a steady stream of interviews being conducted throughout the 52 weeks of the years. In this way, the data can be shown in the form of weekly, monthly data, and accumulated in the form of moving totals on the key metrics, such that the peaks and valleys of the campaign's effectiveness can be matched to the actual points in time of spending. The typical sample size in an Ad Tracking project might be an annual sample of 1000 to 3000 interviews, again depending on such issues as anticipated media weight and likely competitive activity.

The Key Quantitative Metrics

Attention

Because it is transient, attention is difficult to measure. Attention can be inferred by asking the consumer to indicate which visual elements from an ad he recognizes after the ad is removed from view. To measure attention by recognition requires a system for selection of the ad content and a means of interpreting the pattern of responses obtained. A simpler means is that of passively observing where a consumer's eyes are directed when she is exposed to advertising by tracking with a video camera and a reflected light beam. The measure of attention involves both how much time is devoted to a particular element of an ad, and the sequence in which an ad's elements are viewed. When measures of eye movement are combined with specific questions of consumers on the more-traditional measures, an effective and insightful analysis can result. (Weinblatt, 1999)

Recall of Brands and Sales Points

In order for a given advertisement to be effective, consumers in the target audience, such as buyers of the given product category, need to not only have their televisions tuned to the station airing a given commercial, but they need to see the commercial, and to connect the commercial to the brand being advertised. The visuals in a commercial may easily register with consumers, and can sometimes even entertain consumers. But if there is no registration, in the consumer's mind of the brand being advertised, millions of dollars of the client's advertising budget can be wasted.

Because of this fact, advertisers have been interested in determining the extent to which the consumer is able to accurately recall the brand, i.e., connect the advertising with the brand. Over the years, many methods of measuring brand recall have been developed. And several companies have emerged with varying approaches to this measurement.¹

There are interactions between the media and the message communicated. For example, a series of ads all run concurrently on television are called a pod. Experimental data suggests that ads

placed towards the end of a pod of commercials may generate higher brand and sales message recall. However, other data suggest that radio commercials may benefit from placements towards the beginning of a pod.

One of the central issues with respect to the measurement of recall involves the approaches used by the research companies to simulate the experience in which consumers are actually exposed to advertising in their homes. Some companies may recruit consumers and then expose them to the test advertising in a central location, such as a mall, or in a hotel ballroom. Others may send videotapes to consumers at home. And still others may ask consumers to look at downloaded advertising over the Internet.

Some companies may expose consumers to a “clutter reel” of commercials, including a series of test and non-test commercials. Some companies may also embed commercials in an actual pilot program, while others may not feel that this level of simulation is necessary.

In addition, the questions used to measure brand name registration vary from company to company, and from media to media.

Some companies may ask the consumer to name the brand advertised, once she says “yes” to a question about having seen any advertising for a brand in a given product category. Others may provide a list of brands to consumers and ask them to name the advertised brand. Others may show advertising to consumers with the brand name omitted, and ask consumers to identify the brand. And still others may prepare a paragraph describing the advertising’s visuals, and ask for the consumer to name the advertised brand. Put another way, some companies use methods of unaided ad “recall”, while others prefer methods more-closely related to ad “recognition.”

The industry does not have an agreed-upon point of view about which method or methods of measuring brand name recall is superior. Nevertheless, there is general agreement that forging a connection between the advertising and the consumer's ability to connect the advertising with the brand being advertised is a critical component to effective advertising.

The communication of advertising content, or "sales points," is another common area of measurement. Here, however, there is less industry agreement about the necessity, or even feasibility, of its measurement. For example, while many advertising campaigns have as their objectives the communication of specific and rational claims or content, other campaigns may have as their objectives the communication of more subtle or purely imagery-related communication. The creation of an emotional connection between the consumer and the brand may be the objective, rather than the communication of specifically-delineated claims. Sales might be maximized by a more-rational ad campaign for one brand, while a different brand might well be better-served by an emotional campaign, containing little if any real product claims or supporting evidence.

Message recall is also measured using a variety of questions, and no industry consensus exists with respect to the proper questions to ask. Still, when 2 campaigns are equivalent to one another on other metrics, and when a rational approach is deemed appropriate, there is general agreement that the campaign which does a better job of communicating relevant sales messages is more likely to achieve its sales objectives.

Persuasion

Persuasion refers to the ad campaign's ability to increase the likelihood of the advertising to motivate target consumers to actually purchase the brand being advertised. It is often regarded as the other most-important measure of advertising effectiveness, along with brand name recall.

Although most copy testing research companies measure persuasion in one way or another, this is an area of intensive R&D, and considerable controversy.

An oversimplified summary of the long-standing philosophical debate on the relative merits of recall vs. persuasion goes something like this:

- What good is persuasion if no one can remember your commercial, or what brand was being advertised?, vs.
- What good is recall, if no one is interested in buying the brand, even when they remember your ad?

Some research companies believe that persuasion is a far more important measure of ad effectiveness, while others place more weight on recall-based measures.

As a general rule, recall is more likely to be important for new products and smaller brands, in other words, for those brands at the earlier stages of the AIDA model. And recall is more likely to be a relevant measure for emotional campaigns, or those brands which are less dependent on the communication of rational sales benefits.

Persuasion is perhaps more likely to be a critical measure than recall for larger brands, and for those brands more dependent on clear and easily communicated rational benefits and claims.

The methods by which Persuasion is measured also cover a fairly wide gamut of alternatives. One popular method is the pre/post lottery, which originated with the Schwerin technique. In this method, consumers are asked, prior to their exposure to any advertising, to select a brand from a list of alternatives within the product category. They are told that they will then receive a package of that brand in a basket of incentives for their participation. After ad exposure, they are again asked to

select a brand from the same list. The pre to post shift in preferences toward the advertised brand is the persuasion measure.

Other methods include a 5 point purchase Interest question across brands, again administered before and after ad exposure. In this case, the Purchase Interest question ranges from whether the respondent will “definitely buy” the brand, to “probably will buy”, to “might or might not buy”, to “probably will not buy,” to “definitely will not buy” the brand. In the case of a new product, Purchase Interest might only be administered following ad exposure, not both pre and post exposure.

A third alternative is a pre and post Constant Sum Question, in which consumers are asked to “divide your next 5 purchases” in this category across the brands in the category (NOTE: Another variant is the “next 10 purchase” Constant Sum.) One advantage of the Constant Sum question is the information it provides about brand switching, since an increase in purchasing for the test brand will be matched by a corresponding decrease in the purchase of another brand.

Emotional Response Measures

As noted in the first section, the use of verbal questioning to study emotional responses introduces task-oriented thought processes, and that has been criticized by creatives and researchers as interfering with the feelings being measured, or seeking to have things described that may not be accessible, or at a conscious level.

Researchers measuring the effects of advertising through physiological measures believe that they are obtaining a more valid measure because the measurement is totally passive – that is it requires no conscious activity of the respondent, and because the measures are made immediately

and continuously throughout the course of the ad exposure. Physiological measurements which are currently being applied to advertising include:

- GSR or Galvanic Skin Response
- Heart rate and/or blood pressure changes
- EEG, or brain wave patterns
- fMRI, functional magnetic resonance imaging
- Facial EMG, measures of the muscles used in smiling and frowning

Non-physiological measures which are used to measure attention, feelings, etc., concurrently with the ad exposure include rotating a dial or moving a computer mouse along a prescribed line to register the magnitude of the selected feeling state under study.

Researchers seeking to avoid verbal interference have employed pictorial scales such as happy and sad faces. These measures are often made after the ad exposure, and can be inquired with regard to the ad, some part of the ad, or the brand itself.

Diagnostics, Content Analysis and Integration

The basic purposes of quantitative advertising tests, prior to airing, are threefold:

1. to decide which of multiple advertising executions is superior,
2. to determine whether or not the advertising, thus selected, is better or worse than the advertising already being run, and
3. to make any necessary improvements in the advertising, in order to optimize its effectiveness, prior to its airing

In order to address the third objective, research companies employ a variety of methods and measures. Some use standardized batteries of diagnostic questions. They then compare the

responses, for a given ad, against their own, or even client-specific, normative databases of the answers to these diagnostic questions. These comparisons then provide insights into the ad's strengths and weaknesses, which can then be turned into specific recommendations on areas in the ad which could be improved.

Another approach is to analyze the content of the advertising vis-a-vis its component elements, such as:

- The number of brand name mentions
- The length of time between the beginning of the commercial and the first brand name mention
- The number of visual package shots of the brand
- The inclusion of a brand logo or symbol
- The number of times such a logo or symbol is shown
- The number of scenes included
- The number of characters
- Whether the ad is in the style of a "lecture", in which the message is communicated directly, via voiceover or character, to the viewer, or a "drama", in which case the commercial is a mini-vignette, to be watched by the viewer
- The inclusion of a demonstration of product superiority
- The inclusion of music
- Type of music
- Inclusion of a jingle
- Presence of children/babies

Based on research company-sponsored R&D, evidence exists of the connection between the ad's scores on key measures, and advertising content. For example, the literature suggests that there is a systematic relationship between the number of brand name mentions, and the ability of consumers to recall the brand's name, and even that the mention of the brand in the first several seconds of the commercial will tend to aid recall. In this way, an analysis of the ad content can aid in diagnosing the ad's ability to generate high scores on recall, persuasion, and/or likeability (Baldinger, 1991; Haley & Baldinger, 1990).

A third area of diagnosis is that of second-by-second measurement of the content of an ad. Consumers can for example, be asked to react to the commercial while simultaneously moving an electronic dial or down, signifying the consumer's specific positive or negative reactions to the particular scene or components of an ad. When this "interest trace" then turns above or below a steady state line, insights can be gained on the specific elements of an ad, including specific actors, tone, or scenes.

Another approach is to use a group of photos of various people of a variety of lifestyles, and then match these photos to either the brand or specific advertising. Insights can thus be gained about whether or not a specific piece of advertising copy is properly communicating the desired imagery.

Sales Effects

The most difficult, and simultaneously the most important, challenge in Ad Testing is the measurement of the sales effects generated by advertising.

There are many reasons for the fact that this phase of measurement is difficult. For example, survey-based copy tests, by their very nature, cannot include a direct measure of actual sales effects. Intent to purchase measures were created to simulate purchase behavior, but there is no statistical

evidence of a direct correspondence between “definitely will buy” intentions, and actual purchase rates. (NOTE: The literature suggests a rule of thumb that roughly 75% to 85% of those who say “definitely will buy” will actually buy the brand, if they become aware of the brand, and if the brand is adequately available where consumers shop. Similarly, somewhere between 10% and 40% of those who say that they will “probably buy” the brand will actually buy it, under ideal conditions. But, there are many intervening factors that cause these percentages to shift upwards or downwards, such as category dynamics, purchase cycles, and competitive activity.)

In addition, there are many factors, beyond advertising alone, that intervene between the exposure of the consumer to advertising, and the actual sale of the brand. Media from multiple outlets complicates measurement, as do product quality changes, pricing changes, promotional activities such as coupons, and the retailers’ own influences on a brand’s sales.

Consequently, while most experimental measurements of the sales effects of advertising involve a real-world exposure of advertising, it is still not a simple task to identify a definitive ad effect.

One system for doing this is the use of in-market split-cables. The best-known existing split cable system is Behaviorscan, which is a service of Information Resources, Inc. In this system, when advertising effectiveness is the test’s objective, every attempt is made to eliminate all test variables, except the advertising execution itself. A typical test might consist of a 2 market test, run for 6 months or more. Small markets are part of the Behaviorscan system, such as Eau Claire, Wisconsin, or Pittsfield, Massachusetts. A brand will run Campaign A, in one of the cable systems in the test market, and Campaign B through the second cable system. These markets were originally chosen based on the fact that the towns had 2 separate and roughly equivalent cable systems, without significant demographic skews in the subscriber base of either cable system.

Advertising sales effects are measured predominantly using a diary panels in a Behaviorscan test. In each market, thousands of respondents have been asked to fill out diaries which track their purchases across multiple Packaged Goods product categories. Two sub-panels are demographically matched to each other, with one panel being exposed to Campaign A, while the other sees Campaign B. The same pattern of actual programs, dayparts, and actual media spending is used for each campaign in the test. (Note: A “daypart” refers to the time of day in which the advertising time has been purchased, such as Prime Time, from 8 to 11 pm, daytime, late night, the “early fringe” or early evening newscast period, etc.).

Sales are also measured via the scanning data, provided by the retail outlets in each market to Behaviorscan. However, it is not possible to disentangle and separate the scanning data by Campaign, since households exposed to both campaigns will shop in a given retail outlet. (NOTE: IRI did an extensive analysis of the effectiveness of advertising, using hundreds of Behaviorscan tests as the database, in the early 1990’s. This was called the “How Advertising Works” project. Interested readers should contact either IRI or the ARF for the published findings of these analyses).

The major drawbacks of using this kind of system for measuring the sales effectiveness of a given ad campaign are the time and expense of performing these tests, as well as the fact that there is a loss of confidentiality in this approach. That is, competitors may be able to read and react to a new campaign, prior to its ability to gain its full national effect, by observing the activities in these well-known test market cities. Consequently, the number of tests which are evaluated via such systems today represents a small minority of all ad campaigns run on air.

Another system used to measure ad effects is simply to run the advertising campaigns on air, and then use modeling approaches to disentangle their sales effects, using a combination of national

scanning data, national electronic diary panel data, and Marketing Mix modeling. These approaches are covered in depth in Dr. Tellis' modeling chapter, Chapter 24, of this Handbook.

Roles and Responsibilities

The Client

Ultimately, the responsibility to measure and test the effectiveness of advertising rests with the company who is paying for it, and who has the strongest vested interest in its success: the Client. In most companies, the people within the advertiser most-involved with this process include the Brand/Product Manager, the Marketing Director/VP, the Advertising Director, if such a function exists, and/or Corporate Top Management.

The decisions that these management people must make include:

- Whether or not to advertise the brand
- How much to spend on the advertising
- How the budget should be allocated by media type, time of year, time of day, and geography (e.g., nationally, vs. individual market “spot buys”)
- Whether the pattern of spending should be flighted or continuous (i.e., whether the spending should be concentrated within a given period, then repeated in another concentrated period weeks later, or simply spread evenly over the course of the year)
- Which Ad Agency to retain for each media type
- Whether or not to shift ad agencies
- Whether or not to change the advertising currently running
- Which new ad to run

- Whether or not to test and/or track the advertising
- How to test the advertising
- And, how to adjust all of the above over time

Most of these decisions are made as part of a Marketing Plan, prepared by the Brand Manager, as approved and adjusted by those in upper management.

The Ad Agency

The successful Ad Agency serves in a strategic partnership role for the Client corporation. Its role is to create successful advertising for the Client. However, its role in the creation and evaluation of advertising has shifted substantially over the course of the last 20 years. While many ad agencies serve a strong role in both the creation AND evaluation of advertising, many of today's Ad Agencies are much more focused on the creation of advertising, than on its evaluation as to effectiveness.

Partly as a consequence of this trend, many ad agency research functions are housed within a Planning Department, rather than Marketing Research. The term reinforces the Agency's role in the upfront creation and planning of advertising.

As covered above, many Ad Agencies serve a central role in the earliest stages of a Copy Exploratory, and will often conduct qualitative or other creatively-oriented research projects designed to generate the most productive avenues for further study. The evaluative phase is, however, often turned over to the Client.

However, many Ad Agencies also believe that they must serve a broader and more strategic role for their clients than simply the creation of revenue-producing advertising.

Some have generated their own databases and theoretical models, which help to describe the distinguishing features of more successful brands, in contrast to those which are likely to be less successful. In other words, some Ad Agencies serve a helpful and strategic role in the measurement of Brand Equity. This has been one way in which Ad Agencies have historically distinguished themselves from other agencies. Many of the larger Ad Agencies have developed proprietary models and approaches, all of which serve to aid their clients in arriving at an improved understanding of their Brands. Young and Rubicam's BAV, or Brand Asset Valuator model, is one of many such approaches to the issue of arriving at an improved understanding of brands.

Many Ad Agencies also include Media Planning as an active function, which serves the role of intermediary between the Client and the Media itself. In other words, the agency's Media Planning Department will offer detailed recommendations to the Client, concerning how the specific media budget for a given brand should be allocated, across networks, geographies, time of year, time of day, specific program types and programs.

The Research Company

Many research companies specialize in various aspects of advertising research. Invariably, they serve as subcontractors, to either the Client, or the Ad Agency. Often, a given research company will attempt to serve as a long-term strategic partner to the Client organization, in a similar manner to that of the Ad Agency. In some cases, this means that the Research Company will introduce multiple services covering the testing of advertising at multiple stages of its development, evaluation and tracking. Some research companies will also take an active role in the building and maintenance of client-specific normative databases of tests and scores, so that each new ad test can be usefully compared to historical scores. Such databases also can assist the research company in

searching for generalizable truths connected to advertising's likely effectiveness, and how to improve it.

The Media

The Media are the link between the Client's advertising and the consumer. Some media concentrate on a given media type, such as TV Networks, vs. Publishing Companies, or Radio Networks.

Since advertising revenue is the primary source of revenue for Media Companies, they have an intense interest in the various Media Research projects which generate the Ratings or circulation figures, which drive the prices the Media can obtain for their programming.

On the other hand, the media rarely take an active role at the earlier stages of the advertising development and testing process.

Technique Variations by Media Type

The central questions in designing and analyzing an effective experimental test of advertising are as follows:

- How to simulate the environment in which consumers are actually exposed to advertising in a given media
- How many exposures to the ad best relate to the real world
- Whether or not multiple test ads can be evaluated within the same test session,
- Whether or not real-world commercial avoidance, via such technology as remote controls, and/or TIVO, needs to be included in the test, and
- The relative analytic weight to place on measures of recall, persuasion, and diagnostic areas of inquiry

TV Commercial Testing

Research companies involved in the testing of television advertising use a variety of techniques to simulate the real-world exposure to television and its advertising. Many of these methods have been kept the same for decades by the individual research company. Since there is considerable variation in the methods used, and since no single research company has come to dominate Ad Testing , it is safe to suggest that no single method of simulating exposure has emerged as being demonstrably superior to another.

For example, some research companies believe that in-home exposure is important. A method to accomplish this is to send VCR tapes to members of pre-recruited mail panels. After the consumer watches the tape, which includes a pilot TV program, and in which test ads have been embedded, a questionnaire which includes recall questions is administered, perhaps 24 hours after the tape is viewed. Then the consumer is asked to continue watching the tape, and is asked pre/post persuasion and diagnostics after the second test ad exposure.

Other research companies recruit respondents into a central location, such as in a mall location or a hotel ballroom. These techniques often include a process of embedding the test ads within pilot programs, and attempt to simulate real world exposure to the ads by implying that the pilot is the central stimulus for research, rather than the ads placed within the program. Again, persuasion and diagnostic batteries are often administered after a second exposure to the test ad or ads, within the same tape. Most such test systems include multiple test ads within the same test sessions, always from non-competing categories, and often for multiple clients. In this way, the expense for running the session is spread across multiple clients, which keeps per project costs down, and/or increases the profitability of sessions for the research company.

Another variation is the use of a remote control device, to allow the consumer to switch channels during the session. In this variation, the test ads are contained in all programs, which is called “roadblocking” the ads. This can provide useful diagnostic insights concerning the likeability and probability of generating recall for an ad.

Radio Ad Testing

An interesting way to simulate the exposure to Radio ads is to recruit respondents into a central location, and then invite them to watch a video simulating a driving experience. A radio plays in the background, which contains the test ads. The consumer is then asked about recall of the radio ads, followed by a re-exposure and diagnostic questionnaire.

A variation is to allow the changing of the radio stations, this providing a measure of the radio ad’s likeability.

Print Ad Testing

Many of the research companies with businesses in TV-oriented copy testing and tracking also offer services in print testing. This is also an area which seems to be shifting rapidly in the direction of Internet-based data collection, since the exposure of a print ad on a computer screen offers a natural alternative to traditional methods, which are often slower to administer, and far more expensive than might be the case using the Internet.

As in other areas of Ad Testing , there are a variety of methods by which consumers are exposed to print advertising. In some cases, a test ad is inserted, or “tipped in” to a normal magazine issue. Consumers are pre-recruited in a central location or mall, and are asked to read the

magazine in as normal a manner as possible. They are then asked to name the ads recalled. Consumers are then re-exposed to the test ad and are asked diagnostic questions.

A variation on this method is to recruit respondents by telephone, then mail a magazine to respondents, in which the test ads have been inserted. Consumers are then called by telephone the day following the magazine's receipt. They are then asked for ad recall, are asked to look at the test ad or ads again, and then complete the diagnostic interview, including a pre/post persuasion measure.

Another approach is the Starch Readership Starch. Starch was one of the originators of large scale studies of ad readership and effectiveness, and still conducts large numbers of print ad test a year. Starch measures 3 levels print ad attentiveness and readership:

1. Read Most – the extent to which consumers read half or more of the ad's body copy
2. Associated Reader – the extent to which consumers noticed the ad and read enough to recall the brand name, and
3. Noted Reader – the extent to which consumers noticed the ad but did not read the copy

Scores in each area are then compared to Starch's normative database, which can provide valuable insights into an ad's performance. The service is also able to provide insights into the relative strengths and weaknesses of the ad's headline, photos used, color vs. black and white, and specific body copy.

Print Ad Testing is another area in which eye movement measures can prove valuable, inasmuch as insights can be gained about which particular visuals or word are the most likely to grab the consumer's attention, the relative length of time that will be spent on various elements, and the order in which the ad's elements are likely to be viewed.

Newspaper Ads

There are 2 kinds of newspaper ads: ads that contain ad content, and ads that are simply alerting respondents to a weekly special price of a given brand or item in a local retail outlet. The second kind of ad is rarely tested in advance for its effectiveness, but may be evaluated as part of a larger promotional campaign taking place between the manufacturer and the retailer.

The content-oriented newspaper advertising follows the same basic elements as a print ad, and when tested, would follow many of the same procedures as are employed in print tests. However, many of these ads are run on a transitory basis are changed constantly, and are run only on a local or regional basis. For many of these reasons, newspaper advertising is not evaluated with the same frequency as is the case with the other forms of media already mentioned.

Billboard Ads

For many of the same reasons as were just covered in newspaper ads, billboard ads are not always tested for their effectiveness, prior to being placed.

However, when a national campaign is contemplated, and large amounts of media are being expended, or when the billboard campaign is being integrated thematically with a simultaneous campaign being run in other media, the wise advertiser will do a pre-test of his billboard campaign.

In this case, a simulation of an exposure to a test billboard can take the place of a video, in which the view outside a moving car can be simulated, and the test billboard ad can be inserted into the videotape. Such tests are normally conducted within a central location, and/or mall environment. The video may contain several ads, some of which are standard “control” ads, and some of which may be test ads. After viewing the ads, consumers are again asked for ad recall. A

re-exposure to the advertising can then take place, and an interview can then take place, on diagnostic measures and persuasion.

This is another media in which eye movement can provide useful insights into ad content exposure.

Online Ad Testing

Online advertising is still changing rapidly and evolving. Many dollars are spent by advertisers on banner ads. In this case direct measures of effectiveness can sometimes become available, via click-through rates.

However, the early literature on click-through rates indicated that many banner ads were generating low levels of response, with average click-through running at one half of one percent of consumers exposed the web page, or less. Also, many banner ads were so short and fleeting that little in the way of traditional ad content could be realistically included.

Pop-up ads are one way in which online ads can increase their intrusiveness and click-through rates. However, since such ads are often met with annoyance by consumers, technology is rapidly emerging to prevent pop-up ads.

Online advertising is the youngest of the media, and no single Ad Testing service has yet emerged with a dominant place in the spectrum of testing systems.

Innovation and Future Challenges

The Internet

The exposure of most forms of advertising to consumers, using the Internet as an exposure mechanism, is likely to revolutionize the current methods by which ads are created, evaluated and tracked for effectiveness over time. Many of the research companies involved in Ad Testing are either rapidly adapting their systems to the Internet as a replacement for traditional methods of performing advertising-related research (Spielman & Klein, 2001).

The reasons for this are reasonably clear. It is very likely that internet-based ad research will be both faster and less-expensive than performing such experiments using traditional methods. Large amounts of costs can be saved by removing the interviewer from the process of performing ad tests, and by simply replacing the interviewer with a self-administered interview in-home. The time and expense necessary to prepare advertising stimuli, send them to facilities around the country, or to mail paper questionnaires, are areas in which the electronic transfer of ad stimuli, questionnaires and data, can all save large amounts of time and money. In fact, most of this technological transformation has already taken place.

One of the only remaining obstacles to this transfer of methods is the extent to which online samples are representative of the population being measured. While Online Penetration has, for example risen to the range of 60% to 70% of all U.S. households, there are still questions about the demographic differences between internet-enabled households and those not yet online. The question of whether or not the demographic weighting of household would solve all of the possible bias problems is as yet not-completely resolved.

But another, and perhaps even more-important, question is the extent to which consumers are not only internet-capable, but have fast broadband connections to the Internet, using such technologies as cable or DSL lines. Current estimates are that somewhat less than half of all Internet-capable households have as of yet moved into broadband connections.

This is an important issue because of the time that it takes to download video, as well as video advertising, in an online environment. One solution which research companies are experimenting with, is to do other things, such as administering preliminary parts of the questionnaire, while the downloads are taking place.

Still, this is an issue largely restricted to television ads. Print ads might easily be testable in an online environment, as might even radio ads, or billboards.

Technology and Commercial Avoidance

Several technological factors have made the testing and evaluation of effective advertising far more difficult, over the past decade. These include:

- Televisions with remote controls
- Cable and satellite TV systems
- A dramatic increase in the number of broadcast and cable networks
- An explosion in the number of TV channels available
- A dramatic increase in the number of radio networks and channels available, as well as magazines, and internet Web sites
- TIVO and other time shifting technologies

As a result, it is becoming easier and easier for consumers to avoid exposure to commercials. This has put increasing pressure on the ability of advertising testing systems to adequately replicate a real-world exposure environment. Many companies have huge databases of scores on critical measures, all based on methods developed years ago, which may make the companies resistant to making any dramatic changes in their testing systems. Nevertheless, there are growing needs for methods which incorporate the consumer's reduced attention.

Some companies are now experimenting with new methodological approaches, which attempt to reproduce the consumer's tendencies to shift away from a commercial, either in the middle of the commercial, or upon repeated exposures. But this will likely be an area of considerable innovation in the years to come.

An Integrated Research Approach

As we have seen, there is a dramatic difference between the way media is spent in support of a brand, and the methods by which advertising is developed, tested and tracked for effectiveness over time. The advertiser has a strong interest in an integrated marketing and advertising approach. He would very much like to see his television campaign work in a synergistic fashion with his spending, using the same campaign theme across TV, radio, print, and online.

Yet the process of development, testing and tracking may well involve dozens of separate projects, all using separate methodological approaches, and even separate research companies. Because research companies tend to develop specific specialties in a given media or research technique, it is extremely difficult to find a research approach that measures the totality of the advertiser's spending in support of the brand.

This is compounded by the fact that the testing process must stop well short of the ad's true objective: sales. The measurement of sales, using observational approaches to measure actual in-store behavior, is a completely separate research domain than the experimental survey-based school, where Ad Testing and evaluation reside.

It is this unresolved question, how to find a single multi-phase and integrated approach to the measurement of advertising, which remains the ultimate future challenge.

ENDNOTE

¹It is not the purpose of this Handbook to endorse the services of any given company. Several companies specialize in the measurement of advertising in Television, while others specialize in Print, or Radio, or Billboards. However, several have multiple services, covering different media, and/or both Ad Testing and Ad Tracking. The major companies involved in Ad Testing, and Ad Tracking, include, but are not limited to: Ameritest, Bruzzone Research, Communicus, Diagnostic Research, Gallup and Robinson, GfK Custom Research, Ipsos-ASI, Mapes and Ross, McCollum Spielman Worldwide, Millward Brown, The PreTesting Company, Research Systems Corporation, and Taylor Nelson Sofres .

REFERENCES

- Adams, A. J., & Blair, M. H. (1992). Persuasive advertising and sales accountability: Past experience and forward validation. *Journal of Advertising Research, 41*(March/April).
- Advertising Research Foundation. (2002). *Making better media decisions*. New York: Advertising Research Foundation.
- Baldinger, A. L. (1991). The case for multiple measures: Is one ever enough? Transcript of the ARF Copy Research Workshop, (September 11).
- Batra, R., & Ray, M. (1986). Situational effects of advertising repetition. *Journal of Consumer Research, 12*(4), 432-445.
- Haley, R. I., & Baldinger, A. L. (1991). The ARF copy research validity project. *Journal of Advertising Research, 40*, 114-135.
- Hallward, J. (2005). Emotions are the equivalent of first impressions. *Ipsos Ideas*, (April).
- Lukeman, G. (1995). Advertising's role in managing brand equity: What we know from 179 case studies. *Transcript of the ARF Brand Builder's Workshop*, (February 13-14).
- Puto, C. P., & Wells, W. D. (1984). Informational and transformational advertising: The differential effects of time. In *Advances in Consumer Research* (pp. 638-643), 11. Provo, UT: Association for Consumer Research.
- Ray, M. L. (1974). Consumer initial processing: Definitions, issues, and applications. In G. D. Hughes & M. L. Ray (Eds.), *Buyer/Consumer Information Processing* (pp. 145-56). Chapel Hill, NC: University of North Carolina Press.

- Roberts, K. (1998). We need a new kind of research. In *Proceedings of Annual Conference of the Advertising Research Foundation*, (March). New York: Advertising Research Foundation.
- Rossiter, J. R., Percy, L., & Donovan, R. J. (1991). A better advertising grid. *Journal of Advertising Research*, 31, 11-21.
- Spielman, H. M., & Klein, A. (2001). Going online: How one research firm adapted to the Internet. *Quirk's Marketing Research Review*, (March)
- Walker, D. (2000). Are we looking in the right places: Pre-testing and sales validation? *Admap Publications*, (February).
- Weinblatt, L. (1999). The death of print has been greatly exaggerated: Print/TV synergy, The efficient way to advertise today. *Experts Report on Print Research*, ARF Week of Workshops, (October).
- Wells, W. D. (1989). Lectures and dramas. In Patricia Cafferata and Alice M. Tybout (Eds.), *Cognitive and Affective Responses to Advertising* (pp. 13-20), Lexington, MA: Lexington Books.
- Young, C. E. (2001). Researcher as teacher: A heuristic model for pre-testing TV commercials. *Quirk's Marketing Research Review*, (March).
- (2004). *A short history of television copytesting*. Ameritest/CY Research.
- Zaltman, G. (2003). *How customers think: Essential insights into the mind of the market*. Boston, Harvard Business School Press.