

Solutions

A System **For Testing** Advertising

Jeffrey Savitz

Advertising testing is a critical process to insure ads run against the target audience will be effective before large budgets are spent. Of course, the true measurement of the effectiveness of an ad is the degree to which it increases sales. However, this is a difficult quantity to measure given all the other things that occur in the marketplace. Thus, surrogate measures are used.

- ▶ **Breakthrough:** How well does the ad break through the clutter of other ads in the same media?
- ▶ **Communication:** How well does the ad communicate the brand and intended message(s)?
- ▶ **Persuasion:** To what extent does the ad concept convince people to buy the particular brand advertised?

Other important criteria are:

- ▶ **Relevance:** How important is the intended message to the target audience?
- ▶ **Likeability:** How much do people like the ad?
- ▶ **Believability:** How believable is the message in the ad?
- ▶ **Confusion:** Is there anything confusing or difficult to understand in the ad? What?
- ▶ **Appropriateness:** How appropriate is the message for the brand advertised?
- ▶ **Uniqueness:** How unique is the message in the ad?

While many copy testing services pride themselves on having norms relating the average score for these criteria, one must view these norms with caution. If you remember the old Burke Day After Recall test, you may recall that 6% of those contacted “remembered” a TV commercial which never aired. Then, too, ads used to develop norms in a particular category may be based on a limited number of competitors perhaps all from one competitor or all from the client. Indeed, norms may be developed based on ads which failed the copy test and were never run!

On the next page, we have outlined a copy testing system free of the weakness of other systems in terms of normative data. Instead, the system provides norms on all of the above criteria using handpicked client and competitive ads. In this way, the client controls the ads used to develop the norms for all criteria. Finally, because data is collected via the internet, the system is fast and fairly inexpensive.

ESTABLISHING NORMS

The system outlined on the next page, is illustrated using finished television commercials. However, any media can be used as well as any form of communication from concept statements to story boards to animatics to finished commercials.

Of course, the true measurement of the effectiveness of an ad is the degree to which it increases sales.

If you remember the old Burke Day After Recall test, you may recall that 6% of those contacted “remembered” a TV commercial which never aired.

A set of 20 television ads that aired over the past two years would be used to develop norms.

Ideally half would be for the client and the rest for key competitors in the category. A set of five ads from other categories would be gathered. Subsequently, twenty clutter reels would be created each containing one of the test category ads and the five other category ads.

A total of 2000 category users would be interviewed on the internet with one-hundred seeing each clutter reel.

After screening, each participant would be shown one of the twenty clutter reels and asked what categories and brands they remembered being advertised on an unaided and aided basis.

Next, the respondent would be shown the one category ad in the reel a second time. He or she would then be asked a series of questions.

| | |
|------------------------|--|
| Message Communication: | What was the main message the ad was trying to communicate to get you to buy the brand? |
| Breakthrough: | What categories/brands of products and services did you see advertised? |
| Relevance: | How relevant is the message in the ad to you personally? Very relevant 5 4 3 2 1 Not at all relevant |
| Likeability: | How much do you like the ad? Like it a lot 5 4 3 2 1 Don't like it at all |
| Believability: | How believable is the message in the ad? Very believable 5 4 3 2 1 Not at all believable |
| Confusion: | Is there anything in the ad that is confusing? Very confusing 5 4 3 2 1 Not at all confusing (IF SO) What is confusing? |
| Appropriateness: | How appropriate is the message for the brand advertised? Very appropriate 5 4 3 2 1 Not at all appropriate |
| Uniqueness: | How unique is the message in the ad? Very unique 5 4 3 2 1 Not at all unique |
| Persuasion: | Now that you have seen the ad, how have your feelings about buying the brand changed? Much more likely to buy it 5 4 3 2 1 Much less likely to buy it |

Average scores for unaided and total category and brand awareness, percentage with correct message recall and average scores for the other seven questions above are computed three times: once for the client's ads, once for the competitors' ads and once for the combined set. A hypothetical dataset are shown below for the light beer category.

In keeping with the brand colors we will label the ten Coors Light ads in blue and the ten Bud Light ads in red. Below we illustrate the norms for each brand separately in terms of all of the criteria mentioned earlier. Both average rating scores and rankings out of ten are shown.

10 COORS LIGHT ADS

| Rank | Breakthrough | Message Comm. | Persuasion | Relevance | Likeability | Believability | Confusion | Appropriateness | Uniqueness | | | | | | | | | |
|----------------|--------------|---------------|------------|------------|-------------|---------------|-----------|-----------------|------------|------------|-----|------------|-----|------------|-----|------------|-----|------------|
| 1 | C7 | 62% | C7 | 91% | C5 | 4.4 | C5 | 4.3 | C5 | 4.6 | C7 | 4.4 | C1 | 4.9 | C7 | 4.2 | C5 | 3.9 |
| 2 | C3 | 59% | C5 | 90% | C3 | 4.1 | C1 | 4.1 | C3 | 4.4 | C2 | 4.3 | C7 | 4.8 | C2 | 4.1 | C3 | 3.8 |
| 3 | C1 | 53% | C1 | 88% | C7 | 4.0 | C7 | 3.9 | C9 | 4.1 | C3 | 4.1 | C5 | 4.5 | C3 | 3.8 | C9 | 3.6 |
| 4 | C9 | 52% | C9 | 88% | C9 | 3.9 | C3 | 3.7 | C7 | 3.9 | C5 | 3.9 | C9 | 4.3 | C1 | 3.7 | C7 | 3.4 |
| 5 | C2 | 50% | C6 | 87% | C1 | 3.8 | C9 | 3.4 | C1 | 3.8 | C9 | 3.8 | C2 | 4.0 | C9 | 3.4 | C2 | 3.2 |
| 6 | C8 | 49% | C2 | 85% | C2 | 3.8 | C2 | 3.3 | C2 | 3.6 | C1 | 3.6 | C6 | 3.9 | C6 | 3.3 | C1 | 3.1 |
| 7 | C4 | 46% | C3 | 85% | C4 | 3.7 | C6 | 3.1 | C4 | 3.4 | C10 | 3.5 | C3 | 3.7 | C5 | 3.1 | C4 | 2.9 |
| 8 | C10 | 45% | C4 | 83% | C8 | 3.2 | C10 | 2.9 | C8 | 3.2 | C4 | 3.3 | C8 | 3.5 | C4 | 3.1 | C8 | 2.7 |
| 9 | C5 | 42% | C10 | 81% | C6 | 2.9 | C4 | 2.7 | C6 | 3.1 | C6 | 3.1 | C4 | 3.3 | C10 | 2.9 | C6 | 2.3 |
| 10 | C6 | 38% | C8 | 79% | C10 | 2.6 | C8 | 2.6 | C10 | 2.8 | C8 | 2.9 | C10 | 3.0 | C8 | 2.8 | C10 | 2.2 |
| AVERAGE | | 50% | | 86% | | 3.6 | | 3.4 | | 3.7 | | 3.7 | | 4.0 | | 3.4 | | 3.1 |

10 BUD LIGHT ADS

| Rank | Breakthrough | Message Comm. | Persuasion | Relevance | Likeability | Believability | Confusion | Appropriateness | Uniqueness | | | | | | | | | |
|----------------|--------------|---------------|------------|------------|-------------|---------------|-----------|-----------------|------------|------------|-----|------------|-----|------------|-----|------------|-----|------------|
| 1 | B9 | 59% | B7 | 90% | B9 | 4.6 | B5 | 4.2 | B6 | 4.4 | B7 | 4.4 | B6 | 3.9 | B9 | 4.3 | B8 | 4.9 |
| 2 | B5 | 58% | B8 | 88% | B8 | 4.4 | B8 | 4.1 | B5 | 4.1 | B5 | 4.3 | B3 | 3.8 | B10 | 4.1 | B5 | 4.8 |
| 3 | B6 | 56% | B9 | 87% | B5 | 4.1 | B9 | 3.8 | B8 | 4.0 | B9 | 4.1 | B8 | 3.6 | B8 | 3.9 | B6 | 4.5 |
| 4 | B8 | 53% | B5 | 86% | B6 | 3.9 | B6 | 3.7 | B9 | 3.9 | B6 | 3.9 | B9 | 3.4 | B6 | 3.7 | B9 | 4.3 |
| 5 | B3 | 51% | B6 | 79% | B4 | 3.8 | B1 | 3.4 | B4 | 3.8 | B8 | 3.8 | B2 | 3.2 | B3 | 3.4 | B2 | 4.0 |
| 6 | B1 | 48% | B4 | 78% | B1 | 3.6 | B4 | 3.3 | B1 | 3.8 | B2 | 3.6 | B5 | 3.1 | B5 | 3.3 | B3 | 3.9 |
| 7 | B4 | 46% | B1 | 76% | B3 | 3.4 | B7 | 3.1 | B2 | 3.7 | B4 | 3.5 | B4 | 2.9 | B4 | 3.1 | B10 | 3.7 |
| 8 | B7 | 44% | B3 | 73% | B7 | 3.2 | B10 | 3.1 | B7 | 3.2 | B10 | 3.3 | B7 | 2.7 | B7 | 2.9 | B1 | 3.5 |
| 9 | B2 | 44% | B2 | 72% | B2 | 3.1 | B3 | 2.9 | B3 | 2.9 | B1 | 3.1 | B1 | 2.3 | B1 | 2.7 | B4 | 3.3 |
| 10 | B10 | 42% | B10 | 71% | B10 | 2.8 | B2 | 2.8 | B10 | 2.6 | B3 | 2.9 | B10 | 2.2 | B2 | 2.6 | B7 | 3.0 |
| AVERAGE | | 50% | | 80% | | 3.7 | | 3.4 | | 3.6 | | 3.7 | | 3.1 | | 3.4 | | 4.0 |

20 COMBINED ADS

| | | | | | | | | | | | | | | | | | | |
|----------------|-----|------------|-----|------------|-----|------------|-----|------------|-----|------------|-----|------------|-----|------------|-----|------------|-----|------------|
| 1 | C7 | 62% | C7 | 91% | B9 | 4.6 | C5 | 4.3 | C5 | 4.6 | C7 | 4.4 | C1 | 4.9 | B9 | 4.3 | B8 | 4.9 |
| 2 | C3 | 59% | C5 | 90% | C5 | 4.4 | B5 | 4.2 | C3 | 4.4 | B7 | 4.4 | C7 | 4.8 | C7 | 4.2 | B5 | 4.8 |
| 3 | B9 | 59% | B7 | 90% | B8 | 4.4 | C1 | 4.1 | B6 | 4.4 | C2 | 4.3 | C5 | 4.5 | C2 | 4.1 | B6 | 4.5 |
| 4 | B5 | 58% | C1 | 88% | C3 | 4.1 | B8 | 4.1 | C9 | 4.1 | B5 | 4.3 | C9 | 4.3 | B10 | 4.1 | B9 | 4.3 |
| 5 | B6 | 56% | C9 | 88% | B5 | 4.1 | C7 | 3.9 | B5 | 4.1 | C3 | 4.1 | C2 | 4.0 | B8 | 3.9 | B2 | 4.0 |
| 6 | C1 | 53% | B8 | 88% | C7 | 4.0 | B9 | 3.8 | B8 | 4.0 | B9 | 4.1 | C6 | 3.9 | C3 | 3.8 | C5 | 3.9 |
| 7 | B8 | 53% | C6 | 87% | C9 | 3.9 | C3 | 3.7 | C7 | 3.9 | C5 | 3.9 | B6 | 3.9 | C1 | 3.7 | B3 | 3.9 |
| 8 | C9 | 52% | B9 | 87% | B6 | 3.9 | B6 | 3.7 | B9 | 3.9 | B6 | 3.9 | B3 | 3.8 | B6 | 3.7 | C3 | 3.8 |
| 9 | B3 | 51% | B5 | 86% | C1 | 3.8 | C9 | 3.4 | C1 | 3.8 | C9 | 3.8 | C3 | 3.7 | C9 | 3.4 | B10 | 3.7 |
| 10 | C2 | 50% | C2 | 85% | C2 | 3.8 | B1 | 3.4 | B4 | 3.8 | B8 | 3.8 | B8 | 3.6 | B3 | 3.4 | C9 | 3.6 |
| 11 | C8 | 49% | C3 | 85% | B4 | 3.8 | C2 | 3.3 | B1 | 3.8 | C1 | 3.6 | C8 | 3.5 | C6 | 3.3 | B1 | 3.5 |
| 12 | B1 | 48% | C4 | 83% | C4 | 3.7 | B4 | 3.3 | B2 | 3.7 | B2 | 3.6 | B9 | 3.4 | B5 | 3.3 | C7 | 3.4 |
| 13 | C4 | 46% | C10 | 81% | B1 | 3.6 | C6 | 3.1 | C2 | 3.6 | C10 | 3.5 | C4 | 3.3 | C5 | 3.1 | B4 | 3.3 |
| 14 | B4 | 46% | C8 | 79% | B3 | 3.4 | B7 | 3.1 | C4 | 3.4 | B4 | 3.5 | B2 | 3.2 | C4 | 3.1 | C2 | 3.2 |
| 15 | C10 | 45% | B6 | 79% | C8 | 3.2 | B10 | 3.1 | C8 | 3.2 | C4 | 3.3 | B5 | 3.1 | B4 | 3.1 | C1 | 3.1 |
| 16 | B7 | 44% | B4 | 78% | B7 | 3.2 | C10 | 2.9 | B7 | 3.2 | B10 | 3.3 | C10 | 3.0 | C10 | 2.9 | B7 | 3.0 |
| 17 | B2 | 44% | B1 | 76% | B2 | 3.1 | B3 | 2.9 | C6 | 3.1 | C6 | 3.1 | B4 | 2.9 | B7 | 2.9 | C4 | 2.9 |
| 18 | C5 | 42% | B3 | 73% | C6 | 2.9 | B2 | 2.8 | B3 | 2.9 | B1 | 3.1 | B7 | 2.7 | C8 | 2.8 | C8 | 2.7 |
| 19 | B10 | 42% | B2 | 72% | B10 | 2.8 | C4 | 2.7 | C10 | 2.8 | C8 | 2.9 | B1 | 2.3 | B1 | 2.7 | C6 | 2.3 |
| 20 | C6 | 38% | B10 | 71% | C10 | 2.6 | C8 | 2.6 | B10 | 2.6 | B3 | 2.9 | B10 | 2.2 | B2 | 2.6 | C10 | 2.2 |
| AVERAGE | | 50% | | 83% | | 3.7 | | 3.4 | | 3.7 | | 3.7 | | 3.6 | | 3.4 | | 3.6 |

APPLYING THE SYSTEM

To apply the system we simply place a test in the fourth position of a new clutter reel and have another set of category users evaluate the ad in the same way the twenty sets of 100 category users evaluated the 20 ads used to develop the norms. We recommend a sample size of 200 for testing new ads.

Once the data is collected it is a simple task to determine where the test ad falls relative to both the average score for all criteria as well as where it ranks on each criterion. These comparisons can be done for the client’s ads, the competitors’ ads and the combined set as shown below.

Let’s suppose a test ad scored as shown in the table below. This ad is above average against both Coors Ads and the combined dataset on all measures except for Breakthrough where it ranks 8th among the ten Coors Ads and 15th among the combined Coors and Bud Ads. In addition, it is about average against the Coors and Bud ads on Relevance and Uniqueness.

| | Score | Rank against Coors Ads | Rank against Coors and Bud Ads |
|-----------------------|-------|------------------------|--------------------------------|
| Breakthrough | 45% | 8 | 15 |
| Message Communication | 88% | 3-4 | 4-6 |
| Persuasion | 4.0 | 3 | 6 |
| Relevance | 3.5 | 4-5 | 8-9 |
| Likeability | 4.0 | 3-4 | 6 |
| Believability | 4.0 | 3-4 | 6-7 |
| Confusion | 4.7 | 2-3 | 2-3 |
| Appropriateness | 3.9 | 2-3 | 5 |
| Uniqueness | 3.7 | 2-3 | 9 |

COMMENTS ON NORMS

The norms we develop will be far superior to those available through copy testing services for several reasons. First, we will have norms based on the client’s past efforts. Therefore, we will know how a new ad fares against client ads actually run in the past. A similar argument can be made about knowing how well a new ad performs against competitive efforts. Second, our norms will be based on ads run, not ads tested. Third, our norms are flexible. As the client and competitors run new ads, we can use them in the computation of new norms.

SUMMARY

If you have advertising that needs to be tested, call Savitz Research Solutions. The norms you use will be your own!

The norms we develop will be far superior to those available through copy testing services for several reasons.

Call
Savitz Research Solutions
972-386-4050
www.savitzresearch.com