

5 learning strategies for improving ad productivity

John Kastenholz, Unilever, and **Charles Young**, Ameritest/CY Research, explain five ways to improve your ads before they air

LIKE THE CHILDREN of Lake Wobegon, all advertising ideas are above average – if you believe the spin. But the statistical reality of the bell curve is that only 30 in 100 ads you test will score above average; 40 out of 100 must, at least the first time you test them, score only average.

No business wants to spend money on average advertising. Competitive advantage is gained in the marketplace only when you spend money behind advertising that is superior to your competitor's.

But all too often, because of the tyranny of the calendar and the need to put something on air when an air date arrives, your advertising manager must confront the reality of spending millions of dollars behind a merely average performer.

A commercials factory

Advertising is the business process by which ordinary products are turned into stars. Unilever is a master of this. It markets more than 1,000 consumer brands globally and last year spent \$3.3 billion advertising them. Like the Hollywood dream-factories of yesteryear, Unilever manufactures movies in hundreds – bite-sized 30-second ones – each at a cost in the US rapidly approaching \$400,000. And this is the tip of the iceberg. In the US, the industry rule-of-thumb is that while 10% of the advertising budget is spent producing a commercial, 90% is spent buying air time, which means nearly a \$5 million investment associated with every TV commercial.

How do you manage this expensive but creative and therefore frustratingly unpredictable process to beat the averages? You can bring research discipline to the process by using five different learning strategies (right) to repeatedly achieve productivity gains in ad performance.

1 Test the creative with a valid performance standard

Who would launch a new product into a national marketplace without testing it

among consumers? It only makes sense to apply this logic to advertising, which is simply another product of human ingenuity and craft. One way of making sure you are not spending money behind average advertising is to set a firm hurdle rate that says an ad cannot go on air unless it achieves an above-average score in pre-testing research.

One secret for identifying superior creative performers is making sure you use valid measures – are the measurements you collect actually related to effectiveness in the marketplace? To answer that, we have put considerable effort and expense into getting the research right.

Just last year, for example, we stopped using day-after recall testing in the US, a measure first popularised by P&G in the 1950s and still the most widely-used commercial pre-testing measure in the US. The impetus for this came from our research (1) comparing a sample of 60 commercials tested in three different pre-testing systems. Our results showed that recall systematically rewarded boring, bland and emotionally uninvolved advertising.

Of course, this does not mean that you should stop using hard metrics to evaluate TV advertising. Instead of recall, we use a quantitative measure of attention-getting power – one that does not rely on an outdated and limited concept of how advertising enters the consumer's mind. Modern research shows that human memory is highly complex and involves several different systems in the brain. The type of memory that recall testing taps

into – semantic memory – is not, for example, the system where emotional experience is recorded, which is known as the episodic system. So, like the joke about the drunk looking for his keys under the streetlamp because that's where the light is, we concluded that the reason for recall testing may have been one of looking in the wrong place for evidence of ad effectiveness because that is where the research lights were shining in the past.

By changing our performance standard to reward more experiential advertising, we also improved the process of working with our ad agencies. As an artificial barrier to advertising creativity, recall testing had for years been a source of friction between agency creatives, brand managers and researchers. By aligning the different stakeholders in advertising development with a common mental model of how advertising works, we are trying hard to improve our ability to collaborate effectively with our creative partners.

2 Rehearse the creative in rough form first

What stand-up comedians would take their act directly to a national TV audience like the *Tonight Show* without trying out their jokes in a small nightclub first? Yet many advertisers go directly into final production of expensive commercials without any kind of rehearsal in front of a real audience. As Senge points out in his book on learning organisations (2), 'The almost total absence of meaningful "practice" or "rehearsal" is probably the predominant factor that keeps most management teams from being effective learning units.'

For that reason, we test many ad ideas as roughs. Most commercial concepts can be executed inexpensively as animatics or with borrowed footage from other pieces of film – ripomatics. These versions cost only about a tenth as much as the final film and are usually good enough for testing.

5 Learning strategies

1. Test
2. Rehearse
3. Experiment
4. Diagnose
5. Learn



John Kastenholz is vice president of consumer and market insight, at Unilever home and personal care in North America.



Charles Young is founder and CEO of Ameritest, an international advertising and brand research company.

‘What stand-up comedians would take their act directly to a national TV audience without trying out their jokes in a small nightclub first? Yet many advertisers go directly into final production of expensive commercials without rehearsal in front of a real audience’

Rough production not only makes it cheaper to screen ideas, but can be used to make good ideas even better. Not all creative concepts are like the goddess Athena, who sprang fully formed from the head of Zeus. Sometimes a newborn idea has to come of age before being launched into the world.

In a review of 50 recent pre-tests, Kastenholz *et al* compared the performance of commercials that went straight to final production, because of timing constraints or other reasons, with commercials first tested in animatic form and then fully produced with the creative benefit of research feedback (3). Ads that had been tested as roughs were +28% stronger in terms of attention-getting power than those that skipped rehearsal and went straight to final film. This means that we can spend less than \$50,000 on rough production and research of an ad idea to make approximately a \$1 million improvement in the

audience impact of the final commercial. That is the value of rehearsal.

3 Experiment with alternatives

All babies are beautiful in their parents' eyes. The same is true for creative concepts. To find the best solution to your advertising needs, you must avoid the suboptimal approach of falling in love with your first idea.

One benefit of rough production, therefore, is that it makes it cost-effective to test multiple options. Like new product development, creativity frequently boils down to trial and error. Thomas Edison tried more than 3,000 prototypes before hitting on the right way to make the electric light bulb.

By the laws of statistics, if the first idea you test has a 30% chance of scoring above average, testing two ideas gives you a 51% probability that one of them will score above average, testing three gives you a 66% chance, and four give you a 76% probability that at least one concept will score above average on the first pass through the testing system. (This is based on the binomial distribution from a series of Bernoulli trials.) For a new product launch, for example, where timetables are extremely tight and you do not want to be forced into the position of launching with an average commercial for lack of better alternatives, it makes sense to test multiple advertising concepts to gain favourable odds for success.

4 Optimise the creative with diagnostic insights

Remember, the first time through testing you can expect four out of 10 ideas to score averagely, simply because of the laws of probability. Sometimes, though, the first execution of an idea produces a diamond in the rough.

Our ongoing assumption with pre-testing is that many ideas that make it as far as quantitative testing have the potential

to be winners – but for some flaw in execution that is holding them back. The primary goal of diagnostic research, therefore, is to help identify missed opportunities.

There is a huge gain in productivity if you can rework an average idea to make it above average, rather than throwing it out and starting over. Time as well as money is lost when you only use research as a filter. The value of diagnostic pre-testing research is optimisation.

To begin with, diagnostic research can help the ad team answer the following when confronted with a disappointing test score: Is this a little idea that has been well executed, or a potentially big idea that has some executional flaw holding it back? In business terms, is this idea worth investing additional time and money trying to fix?

For instance, we frequently find commercials that score only averagely on attention but highly on diagnostics such as likeability, originality or entertainment value – factors that are normally strongly predictive of above-average attention-getting power. In those cases, we look for some structural flaw in the flow of the film that might be fixed with a little re-editing. Think of it as equivalent to fixing a grammatical error in a sentence you are writing.

Next, the job of diagnostics is to define the problem as precisely as possible – so that creatives know what to fix. For that reason, we do not just engage in copy-testing: we use both verbal and non-verbal diagnostic techniques to provide insights into how a commercial is performing. For example, Ameritest's Picture Sorts technique allows us to understand how a commercial is working as a piece of film, as detailed in (4). These film-direction diagnostics – what we call the 'Spielberg variables' – are particularly important because, at the end of the day, the actionability of diagnostic research must be found in the editing room.

In a review of the 100-plus ads tested by a major business unit in the past couple ▶

of years, we found 50 executions that, on the first pass through testing, scored only average on key performance measures – as we would expect from the bell curve. Diagnostics, however, suggested that half of these average commercials had untapped potential. So they were re-edited, based on the research.

As a quality control check, one in four were retested. And all but one of the executions improved significantly, from average to above average, on key measures – an 87% success rate for improving performance among these underachievers.

In terms of both time and money, therefore, the contribution to advertising productivity can be substantial if you take an optimisation approach to research. Looking at the 100-plus ads this business unit produced during this period, nearly half of the ads that were approved for airing were re-edited, or optimised, based on research. As shown in Table 1, diagnostic research, therefore, provided us with a way to save a large number of creative ideas that simply needed polishing, while maintaining above-average standards for performance.

5 Learn from the competition

From a benchmarking standpoint, it is not enough for your advertising to outscore the commercials averaged in some historical norm base. The reality is that you have to beat the other guys' ads – right now. Your advertised share-of-mind is not just a function of share-of-voice or adspend; it is also a function of the relative strength of your creative, compared to the competition.

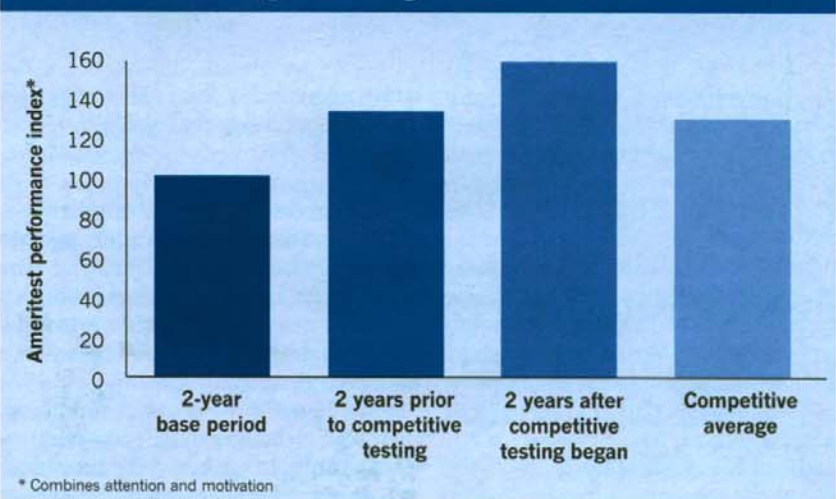
Some years ago, we had a run of weak test scores for a deodorant brand. This led to much discussion and theorising about

Small idea well done	21%
Below average	30%
Above average	28%
Optimised	21%

Note: 117 ads tested
Nearly half the ads that were approved were first optimised.

FIGURE 1

Unilever ads showing a learning effect



deodorants being a low-involvement category where we should not expect test scores to be as high as in, say, the more glamorous shampoos. To some of us, however, this explanation had the look and feel of a self-limiting mind-set.

Then a new marketing director took over, who was a strong believer in research. He reasoned that the competition is selling the same kinds of products to consumers in this so-called low-involvement category, so let's test their ads, to see their scores. Of course, competitors' advertising outscored ours significantly. With this evidence, it was not long before our creative was scoring at this higher level.

Competitive testing enables you to experiment with different ways of reaching the consumer, at someone else's expense. And if you have more powerful diagnostic tools for understanding why consumers respond the way they do to these different approaches, you may end up knowing more about your competitors' advertising than they do themselves. And you learn how to beat them.

The new director was responsible for a number of brands besides deodorants. From the start, he began a programme of systematically testing competitive ads in all his categories. What he learned from this found its way into the performance of his own advertising. When we compared the average performance of our ads

produced during the two years prior to the onset of competitive ad testing to that of ads produced during the two years following, in those same categories, we found that our average had increased by 23%.

Test, rehearse, experiment, diagnose, learn – that is our mantra for improving advertising productivity. Figure 1 shows the performance of 200 Unilever ads tested over a six-year period with this approach. Now a research system that simply filters advertising can raise or lower the hurdle for acceptable performance, but it would not produce a growth curve like this. We believe, therefore, that this improvement in advertising performance over time provides evidence for a genuine learning effect produced by working systematically with these five strategies. ■

1. J Kastenholz, C Young and G Kerr: *Does day-after recall testing produce vanilla advertising?* Admap, June 2004.
2. P Senge: *The fifth discipline: the art and practice of the learning organization.* New York: Doubleday, 1990.
3. J Kastenholz, C Young and T Dubitsky: *Rehearse your creative ideas in rough production to optimize ad effectiveness.* Marketing Research, 2004, under review.
4. C Young and J Kastenholz: *A film director's guide to ad effectiveness.* Admap, September 2003.

john.kastenholz@unilever.com
chuck@ameritest.net