Is Eye-Tracking Making us Blind and Other Research Maladies

Stephen P. Needel, Ph.D.
Managing Partner
Advanced Simulations, LLC.

Marketing Research in a Mobile World
July 2012
Introduction

Marketing Research, as an industry, suffers from a serious malady – low self-esteem. For those of you who pay attention to the marketing research social media, you’ll frequently hear phrases like “Management doesn’t respect me”, “why won’t they listen to me”, and “how did we ever get into this mess”. I usually imagine these are delivered in a whiny teenage voice with the same desperate tone as “I hate school” and “nobody likes me”. Having been in marketing research for a little over 30 years, I can tell you that these were complaints heard in 1980 – so you’d think that 30 years later we’d at least be more innovative in our bitch sessions. It’s good that not all marketing researchers are Buddhists, because if we were all Buddhists, our rate of self-immolation would be alarming.

When I think about this malady, I believe we can approach its solution in two ways – we can use the statistical model or the medical model. The statistical model approach says that if we assume that marketing researchers are normally distributed with respect to skill set, then about half are below average. The solution, from a statistical standpoint, is to get rid of the bottom half of all researchers and we’ll instantly look better to our clients. Now, I prefer the medical model, which tells us to treat the symptoms first, then figure out what caused the problem. So here’s the solution from a medical model point of view:

I am a believer that our fate and our corporate success as researchers are in our own hands, as I have previously written (2006, 2008). If we want to be successful, we need to be the experts that our clients expect us to be. Part of that role means being the gatekeeper or the appraiser or the translator of new technology for your clients. This is what I want to talk about today – using new technology appropriately.

Our industry has developed a slew of new technologies that have the particular advantage of being non-verbal and sometimes unobtrusive; they don’t require us to ask questions of a shopper or a consumer. I want to briefly discuss some of these
methodologies and make the point that these come with their own baggage loaded with lots of hidden assumptions and limitations. I, somewhat facetiously, raise the question of whether eye-tracking is making us blind, whether neuroscience has made us stupid and has facial imaging made us nauseous?

As researchers in a mobile world, we are expected to be faster, more in-touch with the consumer, able to generate insights in a single bound. I’m hoping to slow you down for just a moment and for just a little bit, to take some time to think about what we do and why.

**Eye-Tracking**

Eye-tracking is, today, the most accepted means of measuring attention. The advent of mobile eye-tracking devices along with a push from the neuromarketing and behavioral economics camps, has led to increased deployment of this technology. Eye-tracking was developed in part as a reaction to tachistoscopic presentation of a small set of packages, where recall of a product was the basis for the metrics that differentiated packaging versions. Now, we have the ability to determine what someone focuses on, how long they focus on it, what first attracts their eye-gaze, what brings them back, and so on. This is a great tool, especially as we make the apparatus less and less obtrusive.

However, eye-tracking has a simple limitation - there is no documentation that shows that eye-tracking metrics are related to sales. Many research buyers do not differentiate between attention and sales; they assume that a product that receives more attention will sell better. There’s no surprise here, as some of the purveyors of eye-tracking go out of their way to insinuate this relationship. However, *this is not true*; there is no evidence to suggest that a stimulus that receives more attention will sell more. If we look at the published literature on eye-tracking (cf. Wedel and Peters, 2007 and Chandon et. al, 2009), there is no evidence of a relationship between attention and sales beyond the elementary “if you don’t see it, you can’t buy it”. This is the case in literature put out by academicians and by the companies that sell eye-tracking research.
This doesn’t mean eye-tracking is a bad tool – it only means that it doesn’t tell you what will happen to your sales with one package or another or with one shelf set or another. When an eye-tracking study tells you that viewers are missing a whole section on your web page or that this new package design increases attention to the brand logo or this new version of point-of-sale material grabs your attention faster, it is useful only to the extent that sales are improved by the learning from the study. Taking a marketing action that increases attention alone is a waste of your client’s money. Attention is not necessarily the mechanism that mediates or moderates purchasing. Until that link is drawn via experimental evidence, it is our job as researchers to make that clear to our clients.

**Neuromarketing**

Neuromarketing tries to measure the *assumed* mediators of marketing stimuli, such as attention, arousal, and emotion, by measuring brain activity, with the assumption that more activity is “better”. In their defense, neuroscientists are trying to be more specific about the location of the arousal and its related interpretations, although Damasio’s (1994) oft-cited work suggests that arousal is systemic rather than localized.

Weisberg (2008a, 2008b) makes the point very clearly that people over-estimate the value of neuromarketing data. There is no published, replicable evidence that increased neural activity in humans relates to much of anything when it comes to decisions or preferences. David Penn’s recent article in Research World (2011) makes it clear that the techniques and the data it produces may be interesting, but that neuroscientists have yet to figure out just what it all means. The Advertising Research Foundation’s 2009 review of Innerscope, one of the leaders in the field, finds support for using neuroscience to understand broad emotional response, but they also suggest that the “single-method” approach of FMRI or the EEG headbands are not going to be sufficient, in the same way that in the past GSR or pulse or respiration rate have not lead us to make great strides in understanding physiological responses to stimuli.
Neuromarketers don’t usually tell you that this process is much less scientific than they present. Here are three areas where we can easily be led astray:

- We shouldn’t expect people to have strong neurological responses to most of the products we sell in the CPG world. After all, you are buying a breakfast cereal, not making a life-long commitment! You need to generate a pretty strong response to show up above baseline neurological activity. TV commercials, by their very definition and where neuromarketing started focusing, might generate a strong emotion, but instant mashed potatoes? We don’t think so, and many of the claims made by promoters of neuromarketing are unsubstantiated. Indeed, recent claims by Martin Lindstrom regarding the neuropsychologically compelling nature of the IPhone have drawn criticism in the New York Times from leading neuroscientists (Poldrack, 2011).

- Neurological activity and physiological activity are affect-neutral; you can’t look at an fMRI or an EEG and tell whether someone was reacting positively or negatively to the stimuli. You need some other measure to attach an emotion to a response. In part, this defeats the value of neuromarketing – you still need to ask questions of the respondent.

- When a neuromarketer tells you that there was significant activity in some area of the brain and that means “X”, run away. Neuroscience does not have the ability to identify areas of the brain uniquely associated with choice or preference.

When neuroscience runs the study that shows that, while standing in front of a grocery shelf, we buy the product that most stimulates us neurologically, then we will have a tool that will be useful to researchers studying shopper behavior. However, even when they have tried to do this, it is caveat emptor. A case study with the intriguing and seemingly relevant title “Can Neuromarketing Research Increase Sales?” was published on nielsenwire (Pradeep, 2010). In this study, the researchers are trying to predict the sales impact of one of three magazine covers:
All three designs did well, but the design on the left “ranked highest in terms of overall neurological effectiveness”. So read “ranked” as “not statistically significant” here, otherwise a researcher would say “statistically significant”. Then, the author notes that this cover had sales 12% higher than the same issue as this time last year. There is no test to determine how the other two covers did. Any competent researcher could come up with a number of alternative explanations; here’s a few:

- The more sedate tagline (“Has the fabric of the universe unraveled?”) is more in sync with the magazine’s scientific bent than the bombastic “Torn apart by quantum gravity”, hence better sales – an explanation that has nothing to do with neural activity.
- We have no idea whether a 12% increase in sales is statistically significant – we do not know what the circulation is issue-to-issue and 12% may be within the normal range of variation.
- We don’t know that neural activity discriminated between covers – the article suggests that the winning cover scored exceptionally well in emotional engagement, but was not superior in other primary measures of attention and memory retention.

This is not how you validate a technique.
**Facial Imaging**

A recent Linked-In post leading to an article about Sensory Logic’s Dan Hill (Heller, 2011) touts the use of micro-expressions to measure consumer’s emotional response to products. The rationale is that emotions, rather than logic, drive purchasing, and that we as an industry are not sufficiently adept at eliciting emotions without expensive and time-consuming facial analysis. The writer states that, “the science is complex and the ability to analyze human responses is held by only a few trained individuals.” This, of course, would justify high prices for this research agency and gives it that mystique, that patina of “well it must be good if it’s expensive and time consuming and nobody else can do it” that many research buyers fall for.

There are so many talking points to bring up I hardly know where to start:

- Facial imaging recognition is not a unique and highly specialized research tool – anyone can go on Paul Ekman’s website and get fully trained for about US$39 and he claims to be able to teach you facial recognition in about 40 minutes. Understanding micro-expressions does not have a well-defined protocol, in the sense that parsing emotions into the fine grain one would expect is not done in facial analysis; you are looking at seven basic emotions (anger, contempt, disgust, fear, happiness, sadness, surprise).

- Emotion recognition, like neuromarketing, rests on the assumption that emotions drive our purchasing behavior. This may be true for image-based products or high-end products or emotion-generating products; I’m thinking automobiles, fashion, electronics, and music here. The two problems with this concept are (a) for most of our purchases, which are not in these categories, habit is a much bigger driver than emotion and (b) how emotional do you think we get about our toilet paper or dishwashing detergent? Don’t be swayed by the “emotion drives 95% of purchases” claims that suppliers like to float – unconscious thought drives
about 95% of the purchases in the CPG world (Zaltman, 2003) and there is no evidence that these unconscious thoughts are emotional in nature.

- We might ask, if a researcher can’t measure the difference between these emotions as a reaction to a product or a concept quickly and inexpensively, should they even be in the marketing research business? Do we think that our research participants are lying to us at such a high rate that we need to be constantly on the lookout for deception or that an unobtrusive measure is required? I would contend that if we don’t trust our participants, we shouldn’t be asking them questions. I do recognize that participants aren’t always good at telling us what they are thinking or feeling, but that is a problem with what and how we are asking them rather than they are lying to us. Micro-expressions made for a mediocre television show and have all the promise of being a mediocre marketing research technology.

**Apologies**

I could have focused on any number of other new technologies in this discussion. As an industry, we will waste millions of hours and dollars messing around with Big Data before we rediscover a simple truth – more data is not always better data. We’ve learned this in the past with daily scanner data and single-source viewing and purchasing data and we’ll learn it again with Big Data. The key to making Big Data useful will be twofold – working out the mechanics of integrating multiple data sets and understanding what questions to ask of the data. While we’ve learned some ways to think about Social Media, as you’ll hear from people like Tom Anderson today, what we really know about this today is that it probably doesn’t mean what we think it means. Recent work by Joel Rubinson (2012) suggests that it is not “likes” that matter, it’s time spent at a brand’s site that relates more to purchasing. If true, it bodes ill for driving non-buyers of CPG products with social media. We’ll devote lots of resources to understanding the implications of Behavioral Economics, only to realize that marketing
has assumed most of the principles of Behavioral Economics forever – a lot of what marketers do is get people to make seemingly irrational decisions.

The common them running through the technologies I chose to focus on is unobtrusiveness. As the industry realized that most of us are not very good at asking consumers questions and that consumers aren’t all that good at answering the questions we were asking, we sought other ways to measure and predict their behavior. I’ve always favored the indirect approach to what shoppers are thinking. Analysis of scanner data and household panel data has this indirectness at their core – instead of asking people what they are buying, let’s measure it. Instead of asking people how they think about a category of products, let’s derive it from their purchasing data. My own work in virtual reality takes this same interrogation-avoidance approach. Eye-Tracking, Neuromarketing, and Facial Imaging all assume that the consumer cannot tell us what is going on, so we need to get at their reactions or thought processes another way. I like that aspect of what they are trying to do. What needs to happen is that these methodologies need show they are more than just another piece of data. They need to show that their findings can help marketers increase sales. This they have not done. As researchers, whether you are in a mobile world or a more old-fashioned one, you have to make it clear to your clients this very basic limitation.

Is Mobile the Next Big Thing?

Mobile research needs to prove itself just like RDD (random digit dialing) and Online surveys did during their introduction. A recent study by Pingitore and Seldin (2011) sheds some light on this issue. The authors found that, in comparing Mobile to SMS and Online research, the sample demographics can be similar, and while response rates and drop-out rates were worse, those two problems can be overcome. However, the key finding that should concern us is that the answers they got were different by collection technique. We don’t know if the mobile answers were better or worse, but they were different. If this study holds up, then mobile researchers need to show that they are doing a reasonable job of mapping reality.
**Responsibility**

Market researchers are no less fascinated with bright and shiny things than any other person, and new research techniques are nothing if not bright and shiny. I can safely predict you’ll see a number of intriguing new ideas over the next couple of days; some of them will even be bright and shiny. However, we are in a unique position with respect to these new technologies. It is not our job to pick out the brightest and shiniest. It is our job to make sure that the information we give our clients is properly obtained, that it is reliable and generalizable and unbiased. It is the responsibility of the research suppliers to show that their data is “good”, in the sense that it is a reasonable representation of reality (or at least a part of reality). It is the job of whoever is analyzing the data to use tools that have some generally accepted validity. And finally, it’s the client’s job to test the implications of a piece of research, both experimentally if need be, and against the accumulated knowledge they have about their business.

When someone shows you a new research tool, you should be asking the question of whether it is faster, better, or cheaper than the tools you now have. The rule of thumb is that they need to be two of the three, but that’s not fair to suppliers. Faster is only better if the research buyer can react faster in the marketplace, and that’s pretty rare in most cases. Cheaper is good, but only if it’s cheaper at the same quality – cheap in and of itself is not a desirable attribute of a research tool. Finally, better is always good, as long as better means helping you sell more stuff. A technology that doesn’t help your company or your client sell more stuff is, at best, just bright and shiny. Those researchers who keep on getting a seat at the table do so because they contribute to the company’s bottom line. I’d like to see you all be able to do that – I think you’ll find that helping your clients sell more stuff is, as we Southerners say, the cure for what ails you.
References


