

Target Demographics, Before and After

New Methodologies for Matching Pre-Campaign Audience Plans with Post-Campaign Ad Serving Reports

By RICK E. BRUNER, DIRECTOR OF RESEARCH, DOUBLECLICK
AND KATHRYN KOEGEL, SPECIAL PROJECT CONSULTANT

Executive Summary

Many advertisers want to know the demographic composition of the individuals exposed to their online ad campaigns, along with the reach and frequency of the ads delivered to those target segments. Until now, however, there has been no easy way for advertisers to do this. To address this challenge, DoubleClick commissioned three research companies to devise separate methodologies to produce post-campaign reports for target demographics.

Media research firms comScore Networks, Nielsen//NetRatings, and Interactive Market Systems (IMS) worked respectively with the ad agencies MediaVest, mOne and Universal McCann to produce reports for live campaigns that demonstrated the effectiveness of online ads to access target demographic audience segments within measurable reach/frequency objectives.

The resulting methodologies are relatively easy to execute and can be implemented today using DoubleClick's DART® for Advertisers ad serving platform. This paper describes each methodology and shares findings from their results.

For more DoubleClick insight, go to www.doubleclick.com/knowledge

The Business Need

Most advertisers, particularly brand advertisers, devise their ad campaigns with a target customer in mind. Some audience segments are defined by behaviors (e.g., snowboarders) or by “psychographic” descriptions (e.g., “Urban Achievers” or “Young Digirati”), but the most common method advertisers use to define target audience segments is by demographic characteristics (e.g., age, income, zip code, children in household).

When planning online campaigns, many advertisers rely on audience research panels, notably those of the two market leaders, Media Metrix from comScore Networks and NetView from Nielsen/NetRatings. These panels include tens of thousands of U.S. residents who have agreed to let the research companies monitor everywhere they go online, allowing the research companies to make projections about the demographic composition and usage patterns for thousands of websites. Using those tools, advertisers and agencies can identify which sites and site sections are the best places to reach their target audiences.

After an ad campaign is executed, however, the ad serving reports from a system such as DoubleClick's DART for Advertisers contain only numbers of impressions served, the reach (the unique number of browsers exposed to the ads) and frequency (the number of times individual browsers were exposed to the ads), along with performance metrics such as the click-through rate. Although those numbers all provide valuable information for advertisers, they do not contain

demographic information to indicate how successful the campaign was at reaching its audience targets.

DoubleClick's overall objective in presenting these pilot tests was to demonstrate that there are several effective ways that DART for Advertisers clients can measure online ad exposure on a post-campaign basis today with more precision than is available in any other medium.

(By comparison, TV and print media reporting on advertising effectiveness is based on the “opportunity to see” metric, given that TV viewers may change channels or leave the room when a commercial airs, and magazine readers may never turn to a given page with an advertisement on it. For the online campaigns described in this paper, however, the audience composition reports are based on ads that were quantifiably exposed to users when their browsers displayed specific pages.)

Technical Challenges

There are a few reasons why simply overlaying reports from ad servers on audience panel data is not a straightforward proposition. The two types of systems report on audiences in different ways, and it is important to understand those differences to fully appreciate the strengths and limitations of the pilot methodologies described in this paper.

- Ad server reports reflect a census approach to measuring the audience reached by the campaign; that is, in theory ad servers count every browser exposed to an ad unit. Audience panels

use a projected-sample approach; that is, they include a relatively small subset of the total universe of U.S. Internet users in a panel from which they can make statistically reliable projections about the characteristics of the whole audience.

- Because of their census approach, ad server reports include traffic from international visitors, which in the case of many sites may represent a quarter or more of their total traffic. Audience panels in the U.S. market reflect only the surfing of U.S. audience members. This substantially undercounts the actual number of impressions delivered, but that is acceptable for most domestic advertisers as it is the U.S. audience they most care about reaching.
- Ad server reports count cookie files, which correspond to individual web browsers but contain no identifiable information about those web users, such as demographic characteristics. Therefore, two people sharing one computer could be counted as one person, while one person logging into the same site from two computers (e.g., home and work) would be counted as two people. A certain number of people also block or delete cookies. These cookie issues do not affect the count of the total ad impressions served, but they could skew reach/frequency distributions.
- Audience panels, on the other hand, tend to under-represent the full activity of at-work web surfing, as many large corporations do not allow employees to install the software required by the panel companies. Ad server reports account for browsers in all locations, including at home, work, schools, libraries and so on.
- Audience panels also have difficulty recognizing audience patterns on small sites and small sections within larger sites, where not enough panel

members visit on a regular basis to reliably report on.

In light of those constraints, DoubleClick challenged the three research firms to devise sound, replicable methodologies to enrich ad serving reports with projections of demographic composition of the audiences exposed to campaigns and the unduplicated distribution of reach/frequency to those audience targets.

A more technical explanation of these research methodologies has been submitted for peer review in the Advertising Research Foundation's *Journal of Advertising Research*. What follows are summaries of the three methodologies, along with an analysis of their respective benefits and limitations, in laymen marketer terms.

comScore Networks: Ad Image Capture

Working with the ad agency MediaVest on behalf of a client advertising a skincare product targeted to high-income women aged 35-54, comScore presented a methodology based on image recognition of ad units. MediaVest provided comScore with all the ad creatives associated with the campaign, including graphical and rich media ad types, before the campaign went live in DART for Advertisers.

comScore then used an image-matching technology to recognize all the creative ad units as they appeared in the browsers of members of its 1.5 million-person eXPanded Coverage (XPC) panel. Using this technique, comScore could establish the demographic composition of the audience exposed to the campaign based on their knowledge of their own panel members.

By comparing the demographic composition of the campaign to that of the total XCP panel, which is weighted to reflect the Internet population at large, comScore was able to create an index scale (with a baseline of 100) to show the concentration of target segments relative to the overall online average, as shown in **Figure 1**. For example, women aged 35-44 indexed 205 compared to the average (100), meaning the share of that audience segment in the campaign was just more than double its proportion of the total online population.

In addition, comScore reported the average frequencies at which each of those demographic segments were exposed to ads in the campaign. For example, among women, those aged 45-54 were exposed to the ads most frequently (4.3 times on average), while those under 18 were exposed least frequently (1.7 times). The overall average frequency rate for all demos was 2.9.

Finally, comScore provided the agency with the audience data in the form of an “audience build” trend line. Those charts showed the pattern of how the total audience exposed to the campaign, and the respective target segments, grew over the period of the campaign to achieve its total accumulated reach (known as “cume” in traditional media research lingo).

Demographic Composition of Skincare Product Campaign Indexed High for Women in Target Ages

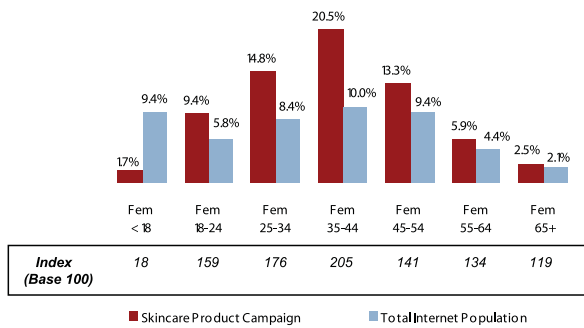


Figure 1 Source: comScore Networks, MediaVest, DoubleClick, 2005

Two-thirds (67.9%) of the people exposed to the campaign were women, an index score of 137 compared to women’s 49.5% share of the overall Internet audience. Women in the household income bracket of \$75,000-99,000 also indexed high, at 135.

comScore also reported the campaign’s demographic compositions for presence of children in the household, education levels, race, and males by age, as well as age groups for both genders combined.

Benefits of This Methodology

The execution of this means of reporting demographic composition was relatively simple for the agency. It just required the agency to send the campaign’s creative units to comScore in advance. From there, the research company was able to identify the ad units when members of its panel were exposed to them and produce the reports without additional involvement from the agency.

This methodology provides demographic composition reports, at the person level, based on an audit of the actual campaign as it was distributed, as opposed to data modeled on historical reports of the audiences for the various sites where the campaign ran. That allows this methodology to produce more accurate, unduplicated reach and frequency numbers

for the total campaign and the various target segments across all the sites on which the campaign ran.

Limitations of This Methodology

The test campaign did not include enough ad impressions for comScore to report the demographic composition on a site-by-site basis. Instead, the demographic compositions that comScore reported refer to the aggregate campaign across all site placements in the campaign schedule. comScore estimates that a campaign would have to include at least 1 million impressions per site in order for the panel to report reliably on demographic composition for individual sites in a campaign.

Another potential limitation of the image recognition technology is its inability at present to identify streaming video in ad creatives, although this is something comScore aims to enable in the future.

Nielsen//NetRatings: Panel Cookie Matching

Working with the agency mOne on behalf of a client advertising a baby product targeting women 18+ with children aged 0-3 in the household, Nielsen//NetRatings used a standard browser cookie approach to measure the demographic composition of the campaign. When mOne prepared the campaign in DART for Advertisers, it attached an extra tracking pixel to each ad unit that corresponded to a cookie Nielsen//NetRatings could match back to members of its MegaPanel of more than 100,000 members. This way, Nielsen//NetRatings could identify which members of its large panel of U.S. Internet users were exposed to the ads in the campaign.

Based on the demographic characteristics it knows about its panel members, including presence of children in the household, Nielsen//NetRatings could then develop reports that showed the demographic composition of those exposed to the campaign. Like comScore, Nielsen//NetRatings expressed that composition with a base-100 index compared to the Internet population at large. It took the additional step to describe demographic segments as “life-stage” clusters, such as “Young Singles” and “Empty Nesters.” Accordingly, it was able to report that the campaign was successful in reaching a high concentration of target segments “New Families” and “Childless Younger Couples,” as shown in **Figure 2**.

Nielsen//NetRatings also reported reach and frequency of exposures for both the

Demographic Composition of Baby Product Campaign Indexed High for New Families and Young Couples

	<u>U.S. Internet</u> <u>Pop.</u>	<u>Campaign</u>	<u>Index</u>
New Families	8%	17%	214
Childless Younger Couples	6%	11%	179
Young Singles	3%	5%	166
Maturing Families	23%	26%	114
Established Families	8%	5%	62
Middle-Aged Singles	8%	6%	82
Middle-Aged Childless Couples	18%	15%	85
Older Singles	6%	4%	73
Empty Nesters	20%	10%	49

Figure 2 Source: Nielsen//NetRatings, mOne, DoubleClick, 2005

“This work provides a good first-step platform for validating the predictive capability of reach/frequency planning models on the web.”

*Gerard Broussard
Senior Partner,
Director of Media
Analytics
mOne, New York*

total audience and the target audience segments, from which it also calculated online gross ratings points (GRPs: the total number of impressions delivered in the campaign, or reach times frequency) and target ratings points (TRPs: reach times frequency for the target segments). The total campaign had a reach of 0.85% of the potential U.S. Internet audience with an average frequency of 3.5, for a total GRP score of 2.96. Of the primary audience target (New Families and Childless Younger Couples), the campaign reached 1.19% of that online audience's potential, with an average frequency of 3.6 for a TRP score of 4.33.

Benefits of This Methodology

Some of the benefits of this methodology are similar to those of the comScore study, although the methodologies differed. The execution of this means of reporting demographic composition was also relatively simple for the agency. During the process of uploading the campaign creatives into DART for Advertisers, the agency had to add one extra tracking pixel per ad unit to match back to Nielsen//NetRatings' panel. Based on that, the research company was able to match back demographic composition to its panel based on the cookie IDs and produce the reports without additional involvement from the agency.

This methodology provides demographic composition reports based on an audit of the actual campaign as it was distributed, as opposed to data modeled on historical reports of the audiences of the various sites where the campaign ran. That

allows this methodology to produce more accurate, unduplicated reach and frequency numbers for the total campaign and the various target segments across all the sites on which the campaign ran.

It should be noted that this approach is not vulnerable to problems of cookie deletion, because Nielsen//NetRatings has software installed on hard drives of its panel members that is capable of renewing cookies as often as necessary.

Limitations of This Methodology

Like the comScore pilot, the test campaign did not include enough ad impressions for Nielsen//NetRatings to report the demographic composition on a site-by-site basis. Instead, the demographic compositions that the firm reported refer to the aggregate campaign across all of the site placements. Nielsen//NetRatings estimates that a campaign would have to include at least 350,000 impressions per site in order for the panel to report reliably on demographic composition for individual sites in a campaign.

Another limitation to Nielsen//NetRatings' approach is that unlike its smaller NetView panel, which media buyers use to plan campaign schedules based on target audiences, the larger MegaPanel used to match cookies in this pilot project identifies only household demographics, not individual person demographics. This means that the post-campaign demographic reports could not specify gender, and age is based on the head of household.

**Interactive Market Systems:
Personal Probability Modeling**

Working with the agency Universal McCann on behalf of four clients (respectively promoting a movie release, a snack food, a consumer electronics product and a business software application), IMS used a data modeling approach to correlate DART for Advertisers ad serving reports to Nielsen//NetRatings NetView demographic panel data.

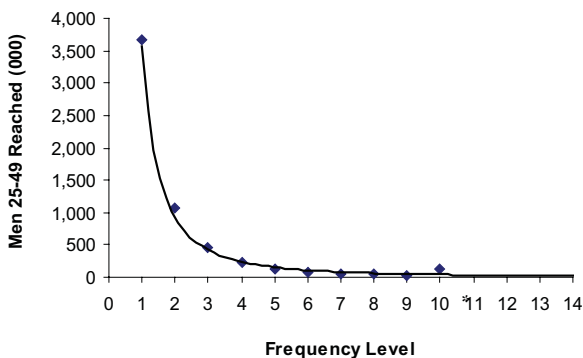
DoubleClick produced reports of the total impressions delivered for each campaign plus the reach and frequency of ads delivered to each site in the schedule of each of the campaigns, according to IMS's specifications. IMS then integrated the weighted data from the DART for Advertisers reports into WebCume, IMS's "personal probability" reach/frequency model for web-based media. ("Personal probability" is a data modeling technique for predicting

audience behaviors that is commonly used in television and print media.) WebCume uses Nielsen//NetRatings NetView audience data to construct its reach/frequency models for individual sites in a given ad campaign schedule.

In this way, IMS could model the demographic composition of the campaigns and the reach and frequency of the ad exposures for each target demographic segment. For example, **Figure 3** shows that more than 3.5 million people in the target audience segment of men 25-49 for a consumer electronics product were exposed to the campaign only once, while only 1.2 million were exposed at the target frequency of 3+, just under 20% of the total target segment that was exposed.

IMS also reported the data in a tabular format for both the campaign totals and target demographics for each site in the campaign schedule, including reach, frequency, GRPs and "reach ceiling" (the maximum percentage of the online target audience reachable for each site).

The Frequency Distribution for This Consumer Electronics Product Campaign Shows That Most of the Target Audience Saw the Ad Only Once



*Frequency levels were grouped as 1 through 9 and 10+; 11-14 as shown in this chart are trended estimates.

Figure 3

Source: IMS, Universal McCann, DoubleClick, 2005

Benefits of This Methodology

One benefit to this methodology for the agencies and advertisers is it requires no additional work on their behalf. DoubleClick transfers campaign data directly to IMS, which generates the reports without the agency having to make any special accommodations, such as providing the campaign creatives in advance, as in the comScore approach, or dropping an extra pixel into the ad server, as with the Nielsen//NetRatings approach.

“Demographic reach/frequency reports allow us to target ad delivery in a more scaleable manner than can be done by solely relying on contextual, behavioral or registration targeting. Additionally, understanding aggregate pressure against a target audience is vital for setting goals and budgets for non-direct response campaigns.”

*Brian Monahan
VP, Group
Communications
Director,
Universal McCann*

Unlike the pilots conducted by comScore and Nielsen//NetRatings, the IMS technique allowed the research company to report the audience composition and reach/frequency distribution of the campaigns on a site-by-site basis, in addition to on a campaign-wide basis. As noted above, comScore and Nielsen// NetRatings should be able to do the same if campaigns include a large enough volume of impressions. By using a modeling approach, however, IMS was able to do so for all four test campaigns in this pilot, with the exception of a few small sites in some of the schedules (discussed in the Limitations section below).

Using IMS's cross-media schedule-combination tool Media Mix, WebCume results can also be integrated with TV, radio and print schedule solutions from other IMS products as well.

Limitations of This Methodology

Unlike the approaches taken by comScore and Nielsen//NetRatings, which found different ways to match back actual exposures of ads in the live campaign to their panels, the IMS methodology models the audience composition of the campaign, based on panel data from earlier months for each of the sites in the campaign schedule. Modeling like this is a well-established technique in traditional media audience measurement, but in the online environment marketers are more accustomed to reporting from actual server transactions. Whether advertisers and agencies will be comfortable with

this more traditional modeling approach for online media remains to be seen.

Some of the smaller sites in the schedules had too little traffic to be included in Nielsen//NetRatings NetView panel, which was the basis of IMS's modeling, so they could not be included in this demographic composition analysis.

Conclusion

As the expression goes, everyone complains about the weather, but no one ever does anything about it. For years, advertisers and agencies have expressed frustration that there is a significant gap between the data they rely on from audience panels for pre-campaign planning on the one hand and from ad servers for post-campaign reporting on the other hand. As described in the Technical Challenges section above, there are fundamental reasons owing to how those two types of systems derive audience data that make it impossible to simply compare the two data sets side by side.

With some determination and ingenuity, however, it is possible to produce sound and practical methodologies to enrich one data set with the other, as these three pilot tests have demonstrated. Reliable reports that detail the audience composition of ad schedules on a post-campaign basis and the reach and frequency of actual ad exposures to target demographics can be provided to advertisers online without great difficulty or technical complexity. Any one of the three methodologies described in this

report can be implemented today using DoubleClick's DART for Advertisers ad serving platform in partnership with the research companies comScore Networks, Nielsen//NetRatings or Interactive Market Systems.

The advertisers that participated in these tests and their agencies generally expressed satisfaction at being able to evaluate the effectiveness of their media buys with these new types of reports.

DoubleClick is pleased to provide market research and pioneer online media

reporting formats that use familiar media research metrics and the vocabulary of traditional media advertising (TV, radio, print), such as demographic targeting, reach and frequency, gross/target rating points, audience build and audience cume. It is our hope that in doing so, we will help traditional advertisers feel more comfortable evaluating the effectiveness of digital media while at the same time helping to educate online marketing specialists in the tried-and-true benefits of traditional media measurement techniques.

Acknowledgements

This study would not have been possible without the help from many talented individuals and organizations. DoubleClick would particularly like to thank the following people and companies for their help with this project: from comScore, Lynn Bolger, EVP of Agency Development and Marie Pauline Mörn, Director of Client Service; from Nielsen//NetRatings, Marc Ryan, Senior Director of Analysis and Manish Bhatia, SVP of Product Marketing and Business Development; from Interactive Market Systems, Craig Gugel, EVP, Worldwide Analytics and Strategy, David Grossbard, Director of Product Development and Medha Patel, research analyst; from mOne, Gerard Broussard, Senior Partner, Director of Media Analytics, Harry Case, Senior Partner, Associate Director of Media Analytics, Jeff Cole, Associate Media Director and Richard Giannicchi, Partner, Group Planning Director; from MediaVest, Adam Gerber, SVP, Director Strategy

and Innovation and James Kiernan, Associate Director of Digital Services and Innovation; from Universal McCann, Brian Monahan, VP, Group Communications Director and Stacy Malone, Interactive Media Director; from DoubleClick, Diane Yu, Director of Engineering, Kurt Ebrahim, Software Engineer and Lynn Tornabene, Director of Marketing for Ad Management. We also thank the six advertiser companies and their staff who worked on this project and allowed us to use their real campaign data anonymously for the sake of this study.

This project was conceived of and managed by Kathryn Koegel, former Director of Research and Industry Development at DoubleClick (currently VP of Marketing at TV Guide Television). This paper was written by Rick E. Bruner, Director of Research at DoubleClick.