# ONLINE MEDIA MEASUREMENT: AFTER THE DELUGE

## Scott C. McDonald, Ph.D., Condé Nast Publications

### Background

The growth of the Internet and the World Wide Web has been chronicled at these symposia since the Berlin meeting in 1995 – by me and by others. In particular, the last symposium in Florence two years ago included several papers on the parallels between online audience measurement and the measurement of other media audiences, on the relationship between print and online, and on the uses of online methods for the conduct of media research. For my part, I presented a detailed critique and comparison of the three emerging approaches to online media measurement: the site-centric, ad-centric and user-centric methods. The present paper will review developments in this field over the past two years. I suspect that much of the ground that I covered in Florence was new to this audience then, but more familiar now. Nevertheless, I shall recapitulate some of the key themes from the 1999 paper as I go.

### No Consensus on Economic Terms for Online Advertising

As I reported in 1999, the online media suffered from a lack of agreement about what should be measured in the first place. While the media sellers wanted to sell their audiences in much the same way that is done for television and print, the buyers of online media pressed for inclusion of advertising response measures in the media evaluation process. In a compromise, the typical online advertising contract of 1999 contained both a guaranteed delivery of a number of impressions (ie. audience delivery) and a back-end adjustment for "click-through" (ie. response guarantee). This represented a hybrid model – conventional advertising crossed with direct-response. To a large extent, this dynamic is still very present in 2001. Online is still viewed by advertisers as a response medium, precisely because of its interactive nature. While online publishers have liked to assert the virtues of the active, engaged, interactive audience, they continue to resist being paid on the basis of click-through. While print and television sellers grouse about being bought on lowest-common-denominator CPM metrics, many online publishers would LOVE to be held accountable only for audience delivery. Despite several efforts through their trade association, the IAB, online publishers have failed to convince advertisers to regard online purely as a "branding" medium; advertisers continue to insist on tracking click-through. Alas, click-through rates for standard banner ads have declined miserably as banners have proliferated – to the ruination of many ad-based sites.

In Florence, I outlined the three ways in which online audiences were being measured in 1999. With some thematic variations, these are still the only available alternatives.

### Site-Centric Measurement

Site-centric measurements are those built up from the website's server log entries. This approach counts the pages requested from the originating web server and thus has the advantage of being a fairly complete census of server activity. This virtue is augmented by other advantages: site-centric measurement captures page requests, regardless of whether the request came from a home user, a school user, a workplace user, from a user in another country, or from a user on a non-computer platform (ie. a PDA, a cell phone, a TV). As was true in 1999, comprehensiveness remains the strong suit for site-centric approaches. However, as I noted in Florence, there are lots of problems with site-centric approaches. They count pages sent, regardless of whether they ever were delivered in full to the end user. They are vulnerable to severe under-counting when pages come from proxy servers or out of local cache on the user's browser. They count pages delivered rather than ads delivered. They can't provide audience demographics. Indeed, they can't always distinguish between human users and robots, spiders, agents or other non-human entities on the Web. They rely on "cookies" to derive unduplicated counts of users and, as such, assume a 1:1 relationship between web browsers and persons that clearly contradicts observed reality. Moreover, the technical mechanics of compiling log entries are idiosyncratic to each website, with pernicious consequences for the validity, reliability and comparability of the resulting estimates.

The ledger of assets and liabilities of the site-centric approach has not changed much over the past two years. However there have been a few developments:

- **Site Registration Campaigns**. Some sites have attempted to put the people back into the site-centric measurement picture (ie. provide demographics) by undertaking ambitious user registration campaigns. Typically, users get some kind of benefit from registration: a customized stock portfolio, free prints from a digital photo, music file swapping, frequent flyer points, and so forth. Large sites like Yahoo! claim to have registered millions of their users. Relying upon cookies, they can go part of the way toward providing a detailed profile of the users of different parts of their site. However because multiple people might use the same computer and browser (or a single individual might have multiple computers with multiple browsers), such registration campaigns still fall far short of an accurate audience count.

- **Harmonization of Methods**. In 1999, I called upon the IAB and other organizations to develop a set of common standards for the compilation of server log data. Alas, progress on this front has been painfully slow. The giddy Dot Com Craze of 1998-2000 provided very little incentive for measurement improvement since the Dot Com media companies were all getting fabulously rich anyway. However by late 2000, the mood had definitely shifted and the IAB began taking steps to address the lack of comparability in site-centric logfile tabulation methods. Early in 2001, the IAB commissioned PriceWaterhouseCoopers to conduct a survey of IAB members on the details of their site-centric tabulation methods and to recommend guidelines for harmonization. The results of that effort are not yet complete.

- **Leading by Example**. One U.S. based site, CNN.com, commissioned an audit of their site-centric and ad-centric methods by the Media Ratings Council, the industry body charged with overseeing the methods of the syndicated research suppliers. This extraordinary step required the MRC board to change its bylaws so as to accommodate an audit, not of a research supplier, but of a media seller. CNN's intentions were twofold: it wanted to use MRC accreditation as a reassurance to advertisers that its site-centric numbers were solid, and it wanted to provide a template or frame of reference to which other site-centric methods would subsequently be compared. The audit process is still in progress.

## Ad-Centric Measurement

Ad-centric measurement was fairly new at the time of the Florence symposium, but its growth path was already evident. It grew out of the tendency of large advertisers to want to serve ads remotely through third-party servers, but developed as the most common currency even for individual media sites. Not all pages on websites carry advertising. Since websites invariably deliver pages of content from one server and units of advertising from another server, the logfiles recording the activity of the ad servers are of more direct relevance to advertisers. As I noted in 1999, ad-centric measurement shares most of the technical assets and liabilities of site-centric measurement in that it is built up from logfile entries. Hence, ad-centric measures typically can't tell you who saw the ad, nor can they estimate unduplicated ad views. Thus, reach and frequency estimates remain elusive. Ad-centric methods are vulnerable to the same sources of over-counting and under-counting as are site-centric methods. None of this has changed in the past two years, but there have been a few other noteworthy developments:

- **Ad-Centric Data Used for Post-Campaign Reconciliation**. As all parties have become more familiar with the reports provided by ad servers, these data have become the most common source of data verifying the delivery of online advertising. These are typically described as "gross impressions" since, for the reasons already described, "net impressions" are almost impossible to derive from ad-centric measurement approaches.

- **Demise of Ad Networks**. In 1999, the online world was abuzz with the growth of third-party ad networks. These offered advertisers the convenience of consolidated ad serving and accounting, but they led to problems with the media sites. The media sites did not appreciate having a third party come between them and their advertisers, and they feared further disintermediation and dilution of advertising revenues. To make matters worse, the ad networks almost invariably reported lower levels of ad delivery than the sites believed to be true from their own logfile records. Before either side could prove the superiority of its accounting methods, the crash hit. One by one, the ad networks fell victim to the capital crunch. AdForce, Engage, AdKnowledge, MatchLogic – all of them bit the proverbial dust. Their demise was aggravated by the high capital costs of their targeting approaches since all of them sold rather complex (and expensive) targeting techniques built on profiles derived from users' clickstream behavior. This had several problems: it was complex and more difficult to understand than old-fashioned content- or interest-targeting; it was expensive to administer; and marketers often found that past behaviors did not foretell propensity to complete a future action. The only remaining ad networks are those, like DoubleClick and 24/7, who also serve as rep firms and service bureaus – and even these networks are sorely reduced in market value and influence. So to some degree, the surviving media sites are in greater control of their ad inventories then they were in 1999, but they are more apt to be using ad-centric measurements to report on their deliveries to advertisers.

## User-Centric Measurement

The user-centric approach is the most familiar to those who grew up in other media since it closely resembles the model used to measure television audiences. Large panels of consumers are recruited, software measures their usage and projects from the sample to a universe estimate. As I noted in Florence, the approach has several advantages over site-centric and ad-centric approaches. It takes the measurement at the point of media delivery, so it is not vulnerable to distortions from caching. It provides audience demographics. It is objective and independent of the media sellers, and it comes in a format that feels familiar to users of other media audience data. However it also has liabilities. It provides little data below the top-level domain; as such, it does not accord with the way that online advertising is bought (usually at the sub-domain level). It has difficulty penetrating non-home environments that are important parts of online usage. And the workings of the meters themselves had not been validated through any independent industry audit. This problem was becoming more obvious in 1999 as the Web developed new, non-HTML elements and as various clever websites tried to devise ways of tricking the meters into inflating their audience counts. So what has changed since 1999?

- **Competition and Consolidation.** The two dominant user-centric measurement companies in the United States, MediaMetrix and NetRatings have survived the crash, though not unscathed. MediaMetrix merged with Jupiter Communications. In 1999, NetRatings (majority owned by Nielsen Media Research) struck a deal for international expansion with A.C. Nielsen; this may have been prophetic since by 2001, Nielsen Media Research and A.C. Nielsen were reunited under the same corporate umbrella again through the acquisitive efforts of VNU. Other players fell by the wayside. PC Data, a company that had attempted to build a syndicated service on a non-random, convenience-recruited panel, went out of business. NetValue, a French-based company that started out using the same dubious panel recruitment approach, withdrew from the U.S. market and converted its European panel recruitment to more respectable and conventional RDD-recruitment.

- **Internationalization**. Both MediaMetrix and NetRatings have both expanded into multiple countries, joined in some places by NetValue. These international extensions are mostly being done through joint-venture arrangements with local research providers, so panel recruitment procedures are not strictly identical across borders.

- **Diversification**. Both MediaMetrix and NetRatings have diversified their offerings to clients. MediaMetrix, through its acquisition of AdRelevance, offers ad tracking data based on monitoring a large sample of sites using spiders. NetRatings offers a similar tracking service, but based on measurements taken directly from their user panels. Because of NetRatings' membership in the Nielsen "family", their tracking service has evolved to conform to the definitions and classifications of Nielsen's MonitorPlus service. In 2001, the NetRatings's ad tracking service also began reporting estimated ad revenues as well.

- **Litigation**. Earlier this year, MediaMetrix filed suit against NetRatings claiming patent infringement because of similarities in the functioning of the two meters. This surprised many observers since MediaMetrix had previously made a big issue of the DIFFERENCES between their operating-system-based meter and NetRatings' browser-based meter. However since some U.S. companies have managed to defend patents for very broad business models or software designs, the MediaMetrix strategy is not to be discounted.

- **Deflation**. The bursting of the dot.com bubble has reduced the amount of attention paid to the monthly traffic rankings published by the user-centric research companies. Though sites still look for opportunities for bragging rights, investors and news media assign this less importance. This is probably a good thing. At least it has diminished the amount of quarrelling about the accuracy of the rankings.

- **Pop-Ups and Pop-Unders**. Pop-up ads were beginning to surface at the time of our last meeting in 1999, but pop-unders are a new phenomenon – or at least they are new to the mainstream. They first started being used at porn sites as a means of extending the teasing of the naïve or timid visitor. Pop-unders are ads that lurk under the current window of your browser; when you try to close the window, they "pop" from underneath. They often can be strung together in long strings of sequential ads. In effect, they cause the user to involuntarily launch a new browser repeatedly. Because the user-centric meters count each browser launch as a site visit, the growth of pop-unders into the mainstream provided a dramatic demonstration earlier this year of how sites could "game" the meters to run up their traffic numbers. The most famous case was that of x10's spring advertising campaign. Using a relentless series of pop-unders, x10 increased its site traffic from January's 232,000 unique visitors to June's 34 million – enough to qualify it for 4$^{th}$ place on the MediaMetrix ranking list. But x10 isn't really a site, just an ad. This provoked a fair amount of tut-tutting of MediaMetrix from the press, the IAB, and even its sister company Jupiter. Rival NetRatings argued for the use of editorial judgment to eliminate x10 from the list since it obviously was not a website in the same way as AOL, Yahoo, MSN, Lycos, Amazon, Disney, eBay, About.com, and the others in the rankings. Indeed, NetRatings made the point that the traffic to the x10 site was involuntary – merely a reflection of people trying to stamp out the ad by closing the window, only to be served yet another pop-under. MediaMetrix defended its ranking by saying that it was reporting behavior, not intentions (which are presumably beyond its measurement capability). Meanwhile, large media sites like Yahoo and CNET, ever more desperate for advertising dollars, weakened their previous "protect the consumer" policies so as to permit more and more of these more intrusive advertising formats to be used on their sites.

- **No Audits Yet**. In 1999, MediaMetrix had begun the process of submitting to an MRC audit, and NetRatings had announced its intention to follow suit. However the process proved to be extremely slow in its development. NetRatings never got through the pre-audit phase, and both firms suspended the process once litigation began in 2001. Thus, formal industry oversight of methods and procedures remains negligible.

- **Workplace Measurement**. Both companies have constituted panels of users at workplaces and report on workplace usage as well as unduplicated home/work usage. MediaMetrix built its work panel by requesting that its home-based panelists also install on their work computers – recognizing that not all would or could do so. This has required relaxation of the customary requirement that all machines be measured if a respondent is to remain in tab; in fact, only a small subset of the at-home panelists have metering software installed on their work computers. NetRatings initially attempted to recruit a workplace sample that is entirely separate from its home-based sample, but found it impossible to develop an integrated, unduplicated home/work audience estimate using this approach. As a result, it established an "overlap" subsample to link the two main panels; this, of course, implies the use of imputation and modeling. Neither company takes a "purist" approach to its workplace panel recruitment, and both admit that the panels are less than perfect. Indeed, there is broad recognition that the problems of workplace panel recruitment are more or less intractable.

All in all, despite some organizational shifts, there has not been enormous change in the user-centric sector over the past two years. The basic assets and liabilities of the methodology remain as they were in 1999. The panel sizes are about the same. The meters operate in much the same way as before, struggling to keep pace with technical innovations in web media formats and elements. There seems to be less complaining about the user-centric services than in 1999, but that may be because users have grown more accustomed to using the user-centric data for estimates of a site's audience reach and composition, rather than as a means of monitoring audience delivery.

Indeed, we can expect to see more scrutiny of website audience delivery going forward. In a paper (previewed at a CASRO meeting in New York last June) to be given in its more fully-developed form at ARF workshops in Chicago this month, Horst Stipp from NBC and Steve Coffey from MediaMetrix analyze online usage according to the same time-based metrics used for TV planning. We have long known that heavy viewers of TV account for a disproportionate percentage of the total time spent watching TV – which is why heavy viewers are discounted by advertisers and light viewers come at something of a premium. The recent work by Stipp and Coffey shows that this is even more true with the online audience. Indeed, their data show that if you break online users into heavy, medium and light terciles according to the amount of time spent online in a month, that the heavy tercile consistently accounts for about 85% of the online minutes! The medium tercile accounts for about 14% of the online minutes, and the light tercile accounts for only 2% of the online minutes. Thus an advertiser inserting an ad onto a website has 85% chance that the ad will be seen by a user in the heavy-using tercile. From this perspective, almost any estimate of a site's unduplicated audience reach for a month is deceptive – since the bulk of the exposures on that site are likely to be to a relatively small set of heavy users.

**Table 1**
**Distribution of Online Home Users by Usage Minutes, October 2000**

| | Users | Minutes | % of Users | % of Minutes |
|---|---|---|---|---|
| Light | 25911 | 1,382,959 | 33.3% | 2.0% |
| Medium | 25856 | 9,198,897 | 33.3% | 13.5% |
| Heavy | 25946 | 57,379,528 | 33.4% | 84.4% |
| Total | 77713 | 67,961,384 | 100.0% | 99.9% |

Source: Unpublished data from MediaMetrix

## New Market Entrants

Against a general backdrop of market retrenchment, collapse and consolidation, there have nevertheless been a few new entrants into the online audience measurement market. Some have exotic names like Red Sheriff , Compete.com, Fovien, and Plurinem. These services have not advanced far enough yet to launch national-level services to compete effectively against the incumbents, and their methodologies are still shrouded in a bit too much mystery to merit much attention here. Some are ISP-based and attempt to look at network-level traffic. Some attempt to combine elements of site-centric and user-centric measurement. All, I suspect, will find it difficult to enter the market at this stage.

There is one significant new market entrant with strong management, aggressive PR, and apparently deep financial backing. In addition, it seems to have developed a mechanism of measurement that is really different from anything we had seen as of 1999. The company is Comscore.com, founded and chaired by Gian Fulgoni, former chairman of the supermarket scanning data behemoth, IRI. Comscore has financial backing from several "heavy hitters" in the Internet and technology sector, including Flatiron Partners, Accel, Siebel Systems and Forrester Research. It also has formed a distribution and sales agreement with a division of DoubleClick called Diameter whereby Diameter will act as a sales rep to further the acceptance of Comscore's syndicated ratings product, NetScore. Last but not least, Comscore has a relationship with NetValue whereby the latter will deploy the Comscore technology in the UK, France and Germany in exchange for the use of the Comscore panel in the U.S. as the source of NetValue audience estimates for the U.S.

In some ways, Comscore offers merely a variant of the user-centric approach. It has recruited an enormous panel of 1.5 million online users through a combination of convenience sampling techniques. It posts recruitment banners all over the Web offering potential panelists one of three different benefits:

- Faster web service & downloads
- Chances to win high-value sweepstakes
- An opportunity to influence the future of the Web

While some banner ads are paid ads, others are placed on a bounty/commission basis through Comscore's affiliate program. Should prospects be lured by one of these three benefits, they are told that they will need to download an applet that will allow their computer use to be monitored, that assures them that their data will remain confidential, and that captures a few demographic variables.

What is different about Comscore's method is the applet that gets downloaded. Rather than act as a meter (a la MediaMetrix or NetRatings), the applet instead re-sets the user's browser settings to redirect traffic to one of Comscore's 200 proxy servers. In effect, every page request from that point forward gets routed through one of the Comscore proxy servers. Because of the size and power of these proxy servers and their closer proximity to the end-users, they supposedly are able to deliver on the promise of enhanced speed (though I have not been able to verify this claim). Though there is no forced turnover in the panel, Comscore reports a 50% natural turnover rate – indication perhaps that the promised speed is hard to deliver. Or perhaps users decide that they don't like having the applet installed after all for other reasons.

Comscore also conducts 1000 weekly telephone surveys using RDD methods to provide a basis for adjusting the weights of the large 1.5 million panel according to demographics and level of online usage. This presumably is supposed to reassure clients that the large panel is "representative", though anyone with a modest amount of statistical training would know that such post-hoc weighting adjustments cannot correct for the selection biases inherent in such self-selected samples as would be generated by Comscore's recruitment methods.

Comscore claims that its measurement method also allows it to track the dollar volume of items purchased online – something that goes beyond the capabilities of either MediaMetrix or NetRatings. With this information, Comscore publishes a "Buying Power Index" that, if believed, would provide a very attractive resource for online media since it would associate online spending even with those sites that do not conduct online commerce. So, for example, even if Vogue.com does not sell online merchandise, one could look at the overall BPI for visitors to the Vogue site and thus presumably make a case for their value to online advertisers. This certainly will spark some interest among online publishers and, indeed, Comscore is now offering 1 year of free "trial" service to members of the Online Publishers Association. Another factor that could endear Comscore to online publishers will be the larger audience size estimates derived from their system. Compared to MediaMetrix or NetRatings, Comscore's numbers are bigger – and in this business there is a tendency always to prefer bigger to smaller numbers (regardless of their provenance).

In effect, Comscore's approach represents something of a hybrid of the user-centric and site-centric approaches. Though users are recruited into a panel, the page request information is not gathered from their computers but rather is compiled from the logfiles of the proxy server to which they have been re-directed. This is rather intriguing, but there are several problems with the approach that should be noted:

- Comscore's large sample affords much greater resolution than other user-centric panels. Thus, reports can be provided at the sub-domain level (though NetScore itself does not automatically do this for any but the largest sites). However the recruitment method would result in a panel biased toward heavy users because it is recruited from banner ads. This means that ALL sites will have higher ratings as a result of the flaws in sample recruitment. Indeed, as already noted, this "flaw" may be tempting to some online publishers.

- Comscore's units of analysis are not people, but machines. It cannot provide an unduplicated estimate of people who use the internet at work and at home, since the person will look like two different machines to Comscore. Similarly, it can't tell the difference between two different people in a household using the same machine. At best, the method can only provide household-level tracking (if one assumes only one computer per household).

- Comscore's applet cannot breach corporate firewalls. Thus, their workplace sample is not likely to be any more representative than those of MediaMetrix or NetRatings in that it will necessarily under-represent employees of large companies.

- Comscore's recruitment promise of faster access would seem likely to result in under-representation of broadband users – the fastest growing segment of the online access market.

- Comscore does not measure activity inside the AOL proprietary system – a significant limitation given AOL's near 50% market share of the online access market in the U.S. To date, only MediaMetrix offers this capability.

- Though it is not alone in this regard, Comscore's measurement method does not advance the art of capturing the non-HTML events that are increasingly prevalent in the online world. Thus, it does not capture the use of Web applications like RealPlayer, Napster, MusicMatch, Instant Messenger, or the like. At present, MediaMetrix continues to provide the broadest coverage of these non-HTML activities, but even MediaMetrix faces a daunting battle keeping up with the constant stream of innovation in Web media elements.

**Summary**

Thus, to sum up: there has not been an enormous amount of change in online media measurement since our last symposium. If you mastered the mysteries of the pros and cons of the three basic approaches – site-centric, ad-centric, and user-centric – you are still on pretty solid ground for understanding the field. There has been some consolidation and retrenchment in the face of the dot.com meltdown. There have been some interesting hijinx with the proliferation of pop-up ads and pop-under ads, a development that has provoked spirited debate about what SHOULD be counted but that also has dramatized the vulnerability of the user-centric panels to forms of "meter gaming". And there has been one major new market entrant, Comscore, with a unique and promising new measurement approach – but with the usual caveats and limitations that undermine its aspirations to tip the field into a wholesale paradigm shift.