BEHAVIORAL METRICS FOR ASSESSMENT OF TABLET AUDIENCES AND ADVERTISING

Dr. Scott McDonald, Condé Nast USA

Following the successful launch of Apple's iPad in Spring of 2010, magazine and newspaper publishers have rushed to create brand-related apps for the emerging tablet platform. Though Apple is expected to continue to dominate the tablet market for the foreseeable future, many other technology companies are racing to launch their own tablet versions. Consumer adoption has been rapid. It is estimated that 60 million tablets will be sold in 2011, up from 17.4 million in 2010. Though app sales figures and surveys indicate that consumers use their tablets for many purposes apart from the reading of magazines, the emerging tablet platform provides a sufficiently attractive target to have tempted most of the major publishers to experiment with the creation of digital editions optimized for tablets. The competitive tracking service iMonitor estimates that by the 3Q 2011, approximately 800 digital editions have been launched by magazine publishers worldwide.

The development of digital magazine editions for tablets has raised questions about the best way to quantify and qualify the audiences drawn to these editions. Though there are a number of variations on the themes, two fundamentally different approaches have been proposed. One approach grows directly from the methods currently used to measure print media: this approach involves retrospective reporting of behavior by samples of claimed readers. The other approach involves the passive monitoring of the usage of the digital editions through the use of tiny "tags" that fire each time a reader takes any action within the digital edition app. Both approaches have their advantages and drawbacks, though there are strong reasons to cheer the benefits and opportunities presented by the passive behavioral approach.

The Self-Reported Survey Approach

Though various efforts have been made over the years to engineer a system for the passive measurement of reading, the printed media have thwarted such efforts. Tests by Arbitron and MRI, among others, based on RFID technology, struggled with numerous problems: reliability of code reception in "noisy" real-world settings, sustaining respondent cooperation, and costs that were never economically palatable to prospective publisher partners. Thus, despite a trend toward passive behavioral measurement for all of the electronic media – and even for out of home media! – print media always seemed to be stuck with retrospective surveys and self-reports as the only viable means of audience measurement. As radio measurement moved from diaries to PPM, continued with the same surveys conducted in home, by mail, or online. As TV measurement moved from diaries to set meters to people meters to audio codes detectible regardless of playback device, innovation in print media measurement consisted of adding survey questions about new sources of copy. The measurement of Internet content and ad exposures primarily relied upon passive behavioral measurement from the outset, with the major debate being about the relative virtues of observing that behavior from the vantage of either the web server or the end user. Only one major service in the US (@Plan) relied upon retrospective reporting. It included several features that were useful for profiling site visitors but the absolute levels of site traffic it reported were never seen as remotely plausible, and last year Nielsen risked significant client upheaval by conforming its audience size estimates to those obtained through passive measurement within its NetViews panel. Virtually without exception, in every medium the marketplace has favored progression to passive behavioral measurement over the vagaries of human recollection and response biases.

As the pace of measurement innovation quickened for the electronic media, many in the print media complained that they were being left behind. Indeed initiatives at the MPA in 2007 and 2008 directly challenged print measurement companies to accelerate the pace of change – with parlous effect on the overall quality of data in the marketplace. Pressed to provide more granular data at an ever-faster pace, measurement companies devised hybrid systems of probability and non-probability samples against which an ever-expanding set of surveys were sent. Increasingly, the non-probability samples were drawn from Internet access panels whose members were explicitly incented to respond to the surveys and implicitly incented to qualify for those surveys – ie. to claim to be readers or to claim to recall transient events like seeing an ad.

Given this history, I cannot say that I have much enthusiasm for the extension of the self-reported survey tool to the measurement of magazine digital editions. The likely flaws stand before us in plain sight. The number of claimed readers will likely be significantly higher than the number of editions in circulation. While some of this surely might be true readers-per-copy for the digital editions, the more likely explanations are the usual suspects: telescoping, replicated reading, confusion about source of copy, and the aforementioned biases of Internet access panels and their incentive structures. At this early stage, we do not yet have enough information from the measurement companies that would allow us to evaluate the plausibility of the digital edition audience estimates derived from self-report surveys, but the smart money would bet that they will be as out of line with the passive measures as was the case with the now-abandoned @Plan method. Recall-based surveys may be able to give us ordinal rankings on how particular editorial features or ads performed in their ability to attract attention or provoke recall – and since recall is beyond the ken of behavior-based systems, that could add important value if we believe the results. And if we think they are reliable and valid enough, recall-based surveys also may be able to give us a unique window into

demographics and buyer behavior. But to earn our confidence in their reliability and validity, they need to demonstrate a consistently close relationship to the behavior-based systems.

Passive Behavioral Metrics

The behavior-based systems for tablet audience measurement closely resemble those used for site-centric web measurement, and indeed the leading purveyor of passive behavioral metrics for tablets is also largest vendor of site-centric web measurement systems (Omniture, now a division of Adobe). Although web and tablet behavioral tracking systems are similar, they also differ in crucial ways.

Web systems require the use of browser-based "cookies" which can be deleted by the user, resulting in over-estimation of unique users by server-based systems. Tablet systems do not use cookies, but rather insert at publication very tiny bits of code associated with each page or data element (interactive overlay, video, slideshow, etc.). The software inserts the codes automatically and the software reporting system will provide consistent rules for the back-end reporting of the usage data; the publisher will simply need to organize naming conventions to make recovery of the usage data simpler. While this can be tedious at first, it can be a very routine and systematic part of the digital edition publishing process. And as with site-centric measurement systems, both the software and the publisher implementation practices can be audited by such third-party entities as the MRC to assure rigorous quality control.

Though passive tablet measurement systems don't have the "cookie problem" that similar web server systems do, they have a different potential blind spot. On the web, the user is by definition always connected to the Internet, since all activity involves calls back and forth between web user and web server. However in the tablet environment, readers download magazine issues to their apps and then can read them without an Internet connection (as, for example, when on a plane or in a subway or out of reach of a Wi-Fi connection). The capture of that "offline" reading activity is a potential problem for passive measurement of tablets. As currently implemented, Adobe's system caches all offline reading activity and reports it back to the measurement system the next time the reader opens up that magazine's tablet app. It would be better, of course, if these cached data went back to the Adobe servers the next time that the tablet got connected to the Internet, regardless of which apps were open. While this is technically possible, it is outside of the current business arrangements between publishers and Apple. Because of this feature, it is possible that some offline tablet reading activity is never captured by the measurement system since, in some cases, readers never return to that magazine's app (or perhaps do so after a sufficiently long interval to no longer be of interest to either publisher or advertiser). As such, some undercounting of activity is inevitable. Based on our monitoring at Condé Nast so far, it appears that about 15% to 20% of reading takes place while "offline", but that is only the view of the activity that was captured by Adobe servers. We have no way of knowing yet how much will end up being reported long after the event (by very infrequent readers) or never at all (by readers who do not return to read another digital edition of that magazine).

Passive tablet measurement systems also differ from web measurement systems in their treatment of the "page" as a unit of measurement. On the web, the user clicks to initiate each page download: though some pages go beyond a screen frame, this is relatively infrequent, so each page view is associated with a discrete server call initiated by a user click. In the tablet environment, the reader navigates by using his or her finger to move from page to page. Sometimes, this navigation is rapid – as when someone retraces pages already seen to move to a different part of the issue. This poses a problem since we have no commonly agreed-upon definition of the length of time one should alight on a page for it to count as having been viewed. In practical terms, the Adobe tags fire whenever a page snaps into its frame, a process that takes about half a second. So we could accept that as a *de facto* page view and use the time counting mechanism to classify views according to their duration. Or we could devise some other edit rule to only qualify pages as "viewed" when the reader alighted for a longer time.

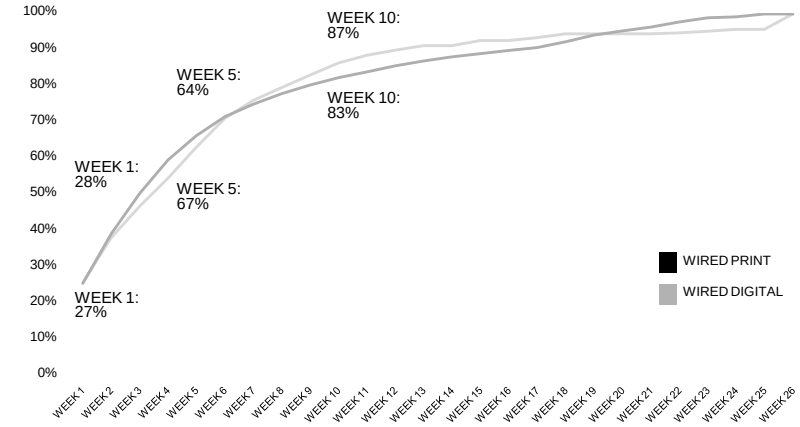
Page definition is also confounded in the tablet environment by the diversity of screen sizes and form factors available. A story that takes 10 pages in a printed magazine might optimally be presented in 14 pages on the 10" iPad, in 18 pages on a 7" inch tablet, and in 21 pages on a 3" smartphone. So how many "pages" is it in our usage reports? This is already a problem during this early phase when Apple's iPad is the dominant device; it will only get more difficult as the devices capable of receiving the digital editions become more diverse. Even now, the page count is affected by whether the reader is viewing the story in portrait mode or in landscape mode, and by whether or not the story was presented in the tablet in Apple's native SDK code or in HTML5 (which allows for longer running pages that flow beyond the screen, resulting in lower "page" counts). Obviously, these problems of page commensurability and comparability could be solved by a universal (or at least widely accepted) character-to-page conversion formula. Here again, no such conventions have yet been devised and adopted by the industry, however these problems do not seem to be insuperable to me. They simply require consideration, discussion, and the formulation of tentative solutions that then get subjected to empirical scrutiny before being revised. Given Adobe's willingness to work with publishers and other stakeholders, their stated intention to submit to MRC audits, and their dominant position in the server-centric measurement space, they are an ideal company to facilitate the emergence of *de facto* standards for technical definitions of units like a "page".

However passive behavioral measurement does have a few fundamental structural problems that will be difficult to solve. The data collected from each tablet bear the fingerprint of the unique device identification number that allows a count of unduplicated users (ie. tablets) accessing the magazine or newspaper. However, as on the web, devices are not identical to

users. An individual reader may access the same content from multiple devices, each with their own unique device ID numbers; in such cases, estimates of unique users will be over-estimated. Conversely, devices might be shared among households, friend networks, work associates, or ultimately even among strangers in public places; in such cases, passive behavioral measurement will under-estimate the true audiences of the digital editions. Our initial surveys on this question among the Condé Nast digital editions suggest that accessing the editions from multiple devices is very rare at this point. However we find that some digital editions are shared on a single device; when this happens there are usually 2 to 3 readers per copy – lower than the levels achieved by printed magazines in public places, but comparable to the within-household levels of sharing estimated by the print audience measurement services.

After a protracted period of negotiation of business terms for the sales of digital subscriptions, Condé Nast and Apple finally signed a deal in May of this year that permitted the sales of all-digital subscriptions on iPads, and that also allowed current print subscribers to gain free access to the digital editions. Almost overnight, a large number of subscribers rushed to "authenticate" their access to digital editions and start reading their magazines on iPads. As wonderful as this development may be, it presents yet another difficulty for passive behavioral measurement since we do not know how these "authenticators" will be allocating their time and attention between the digital editions (for which we have detailed behavioral data) and the printed editions (for which we have zero behavioral data). Assume for the moment that a subscriber has qualified for the digital edition and begun to download each issue. Assume further that she allocates 50% of her time and attention to the digital edition and 50% to the printed edition. With our perfect visibility into her digital activity, we will erroneously attribute to her half of the editorial and ad page impressions that were actually delivered, and half of the time spent. Our early interviews and surveys on this subject suggest that, in fact, time and attention are highly variable and situational – with many readers shifting their allocation of attention from issue to issue, depending on a host of circumstances. This, of course, only makes more difficult the problem of capturing the entirety of the reader's activity with the issue. This will be a central focus of research on digital editions going forward and will provide some interesting challenges to the ingenuity of our methods and our models.

Figure 1.
Audience Accumulation of Wired Magazine
Print vs. Digital Edition 2011



SOURCE: WIRED - MRI FALL 2010 ADCUME & OMNITURE VISITS TO DIGITAL EDITION

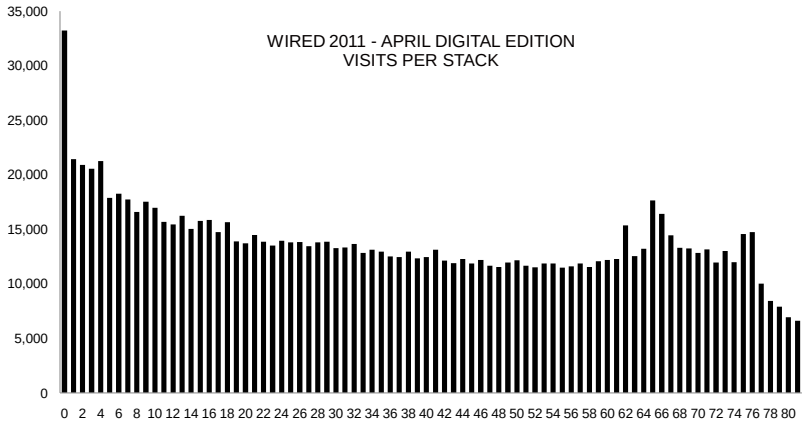
CONDÉ NAST 1

Despite these difficulties, the early data are beginning to provide fascinating evidence that magazine reading on tablets strongly resembles the behavior long observed with printed magazines – and contrasts with the quicker, more fragmented styles of reading associated with websites.

- Though we see a surge in issue downloads and opens as soon as a new issue is published, the accumulation curve for the digital edition of Wired (the only Condé Nast title for which we have a sufficient number of observations at this point) is remarkably similar to the accumulation curve estimated by MRI for the printed edition of that magazine (see Figure 1).
- Readers of digital editions (again using Wired as our example) tend to page through these editions from front to back, with relatively little direct navigation to stories, despite the facilitation of such navigation in the user interface of the

digital editions (see Figure 2). This behavior contrasts sharply with the navigation characteristic of websites – where readers are more likely to jump into the middle of a site and then exit than to start at the top of the site and work their way down its structure in any systematic or orderly fashion.

Figure 2
Page Traffic by Page Location in Issue



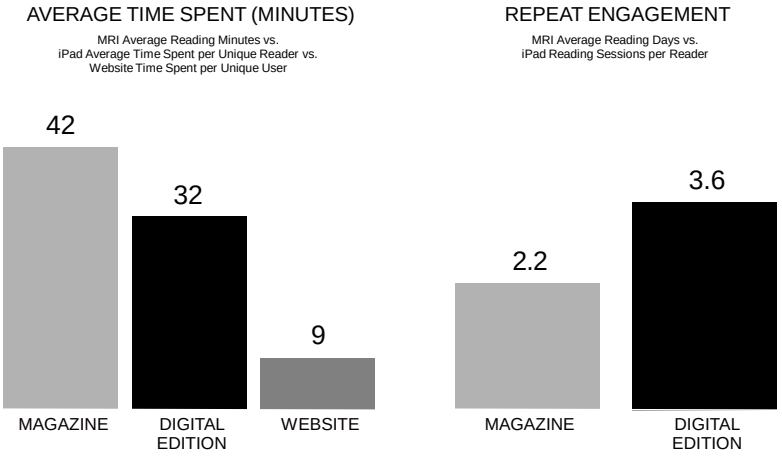
SOURCE: OMNITURE

- Total time spent reading the digital editions, while a bit lower than the estimates made by MRI for the printed editions, is still much longer than the time spent reading on those magazines' websites (see Figure 3). What's more, the average number of sessions per digital edition seems roughly comparable (or perhaps even a bit higher than) the MRI estimates of number of reading days for those titles' printed editions (also Figure 3).

These generalizations surely will change as we get better data and a wider array of titles on which to base the findings; however it does seem that we stand today on the threshold of a very exciting era in magazine media – a time of more transparent visibility into the reading process. By the end of this year, we expect to be providing advertisers with granular counts of the number of discrete "readers" (really tablets) that paused to look at their ads, and the total number of impressions that those ads generated. With a bit more work, we will be able to give them precise information on the ability of those ads to sustain attention across the seconds and minutes of the readers' encounter – time expanses that seem to be particularly impressive when those ads invite and enable direct interaction. And beyond that we may be able to devise new measures of engagement that capture the full range of reading experiences – from casual browsing, looking at pictures, and playing with interactive devices on the page, to old-fashioned immersive reading of long-form text articles. The interaction of these reading styles and reading occasions with measures of ad effectiveness will be fascinating to study.

Of course, self-reported surveys may well retain an important role in our assessments of ad effectiveness by measuring ad recall, ad likeability, brand linkage, consonance with brand communication objectives, and intentions regarding future actions. As skeptical as I am of their current value in this regard, I would welcome having them provide the good solid dependent variables that my imagined future studies to find the best magazine environments for optimal ad placements. Alas, I fear that the race to bring new cross-platform data to market has tarnished studies that once maintained methodological vigor and hastened the degradation of the data that are needed for media planning. While I intend to continue to press the case for more responsible management and use of our venerable self-reported surveys, I think we should all be excited to have, at last, a viable and scalable passive behavioral measurement mechanism for magazine and newspaper media. If the history of the electronic media is any guide to our own future, passive methods will stake a very large claim in the measurement of reading in the 21st century.

Figure 3
Engagement Measures: Print vs. Digital Editions, Spring 2011



SOURCE: SPRING MRI 2011 & OMNITURE; AGGREGATE DATA FROM GLAMOUR, GOLF DIGEST, SELF, & WIRED