Measuring the internet

Measuring the internet

2700 words

Books & Arts

Jock Given

16 August 2012

Digital media users may be easy to track but they can be very hard to follow, writes Jock Given

More media options mean smaller numbers tuned in to any one of them. Nicole Emanuel/The Age

Rating the Audience: The Business of Media
By Mark Balnaves and Tom O'Regan with Ben Goldsmith | Bloomsbury/A&C Black | $45

“THE Net… is different,” wrote Randall Rothenberg in Wired in January 1998. “The Net is accountable. It is knowable. It is the highway leading marketers to their Holy Grail: single-sourcing technology that can definitively tie the information consumers perceive to the purchases they make.”

What we have learned in the decade-and-a-half since is that some of what happens on the internet is much more accountable. Newspapers, for example, have learned which of their stories attract online readers in a way they never knew offline. Digital technology has given us many better tools to measure what people are doing with media.

But it has also enabled people to do new things with media that are incredibly hard to measure reliably. Audiences and users may have become less knowable at precisely the moment when we thought we finally had the technology to pin them down. Some extraordinary machines have been built to drive that highway to the marketers’ Holy Grail, but the road still winds on and uphill.

THERE are several ways we can try to find out what people are doing with media. We can track them; we can ask them; or we can watch them. Tracking them, recording particular actions, seems the most reliable. We know what they did. The old media of live performance and cinema still provide the cleanest examples. People buy tickets; they watch a show. We record who bought the tickets and can safely assume most watched the whole thing, and once only. If they want the media experience over again, they buy another ticket and we record them again. For audience researchers, that’s about as good as it gets.

Other media products also leave behind a recorded transaction – a book, a newspaper, a magazine, a DVD, a paid music download. But while this tells us who bought the product, it doesn’t tell us as much as a cinema or concert ticket about who listened, watched or read, especially if they did it many times.

Online and mobile digital media seemed more like cinema tickets than CDs, a supremely trackable kind of media use. We record the keystrokes and know the user went to that URL, then this one, then sent an email, then checked the weather, Skyped a relative, paid a bill.

At least, someone knows. The sites that users visit can know a lot about their own visitors: how many of them there are, what countries they visited from, where they were online before they came to the site, how long they stayed, and where they went afterwards. Many site owners do this by installing tools like Google Analytics or by getting someone else like Web Trends or Web Analytics to do it for them. What the organisations using these “on-site analytics” don’t discover directly is what is happening on other sites.

That’s what “off-site analytics” can do. These come from internet service providers or panels of users recruited
for the purpose. ISPs can analyse their own data about their users’ activities or provide it to companies like Experian Hitwise, which can in turn supply data to third parties on the behaviour of much larger aggregations of users. Hitwise now says it gathers data from around three million Australian internet users via its ISP partners (though the biggest, Telstra’s BigPond, is not among them), and twenty-five million users worldwide.

Other “off-site” data providers, like Nielsen and ComScore, recruit users to participate in panels chosen to represent the demographic characteristics of the whole user population. The activities of this sample are tracked using software installed on their computers, and aggregated to give an estimate of what all users are doing. By recruiting a particular sample of users, panel measures aim to get closer to the activities of real people with known characteristics, rather than those of whoever happens to be visiting a site or using a particular ISP.

These different methods can be combined. ComScore’s Unified Digital Measurement methodology integrates panel and site data. Since late last year, Nielsen has been using a hybrid approach in Australia, incorporating panel data and information gathered from tracking tags placed by publishers on their own web pages. The Interactive Advertising Bureau of Australia appointed Nielsen as the sole and exclusive preferred supplier of online audience measurement services in Australia in May 2011.

These different methods can tell different stories.

**RATING the Audience** explores the history of the methodologies and conventions governing media measurement from the midst of this contemporary maelstrom. Balnaves, O’Regan and Goldsmith locate the beginnings of systematic audience measurement at the start of radio broadcasting in the 1920s and 30s. Unlike the popular media that preceded it – print, recorded music and cinema – radio listening left no trace beyond the initial decision to buy a receiver and, in some countries, to pay an annual listener licence fee.

The pioneers of radio audience measurement had to come up with a new approach. They decided to ask listeners what they listened to. In an era when most houses had a woman at home during the day, research teams could walk down the street and find, behind most doors they knocked on, someone happy to be asked. Others asked listeners to fill out listening diaries. As more people got telephones, researchers called them up either to ask what they were listening to at the time (“telephone coincidental”) or what they had listened to in the recent past (“telephone recall”).

Developing their methods alongside the pioneers of political polling, these audience researchers had to deal with all the factors that judges in courtrooms are trained to manage. They learned to probe the tricks memories play. Radio listeners might not tell researchers the truth. They might not even know what it was. Did you really listen to the news last night? Or have you forgotten you were late home after a meeting ran over time? Or you were catching up on something else while the dinner cooked? Or chatting about something not nearly as memorable as your normal routine?

Surprising things were discovered. The BBC thought no one dined before eight and was horrified to discover many were finished their evening meals by seven. Audience researchers absorbed these surprises and tested them against their instincts and professionalism and the needs of the parties that wanted their data. Balnaves, O’Regan and Goldsmith attribute “the core of the modern ratings convention” to Archibald Crossley. Hired in 1929 by a group of American radio advertisers, Crossley’s job was to measure the “unseen audience” and develop a mechanism that enabled advertisers to choose which broadcast outlets best reached their target audiences.

Crossley’s system made “exposure” the key measurement. It used a sample of the audience rather than a complete census, produced a “single number” whose “inherent correctness” appealed to all parties, and insisted that any distortion by the ratings provider or subscribers was unacceptable. Crossley’s “telephone recall” ruled in the United States in the 1930s but was overtaken by C.E. Hooper’s “telephone coincidental” system in the early 1940s and then by Arthur C. Nielsen’s “audimeter” later in the decade.

Nielsen was a market researcher who tested new products and determined market shares. Where his predecessors in the ratings game had asked listeners what they did, Nielsen tracked them with an audimeter that picked up and recorded the frequencies radio receivers were tuned to. The information gathered was stored on a wax drum then transferred to film and later solid-state memory. When radio homes started buying television sets
in the 1940s and 50s, Nielsen started measuring their use too. His company became a near monopoly in the new medium, eventually leaving radio measurement altogether, which Arbitron came to dominate.

In Britain, where broadcasting itself was a monopoly, the BBC set up a Listener Research Section, later called BBC Audience Research, which rejected the methods being developed in the United States as “one-dimensional and unreliable.” According to Aberystwyth University’s Sian Nicholas, a “Listening Barometer” was developed to measure “pressure rather than heat.” It used weekly returns from volunteer listeners to measure a variety of radio publics rather than a monolithic radio public.

In Australia, there was competition, deep and sustained, between Bill McNair and George Anderson. Initially working for the advertising agent J. Walter Thompson, McNair published the landmark *Radio Advertising in Australia* in 1937 and built up a ratings business using personal interviews and the recall method. Anderson worked in radio and wanted to know “who was on the other end of the microphone, listening.” In the 1940s, he established a diary system, rejecting the American audimeter as too capital intensive for Australia. The two systems ran in parallel for nearly thirty years (although McNair shifted to diaries in the 1960s), until the two companies merged to form McNair Anderson in 1973.

EMBEDDED in this unusual level of competition in the early decades of Australia’s broadcast ratings, according to the authors of *Rating the Audience*, was “a perception of checks and balances, even if it was costly to run two methods of audience measurement.” Both suppliers thought their methods were best; the competition drove them to be as good as they could be.

Competing ratings systems are attractive because each helps to keep the others honest and innovative. Technical and methodological innovation might throw up different and perhaps more accurate or commercially powerful pictures of audiences, or the scope for more detailed or timely analysis. A single ratings system is attractive, however, because it delivers an industry consensus about audience size and saves money.

With competing providers, the methodological arguments are fought more publicly, day-by-day, as the providers compete for customers. With a single provider, the methodological tussles are more private, bursting out only occasionally in major assaults on the whole system. *Rating the Audience* discusses several of these: the contentious decision to choose Television Audience Measurement ahead of Nielsen when the BBC finally got commercial competition in the 1950s; the United States Congressional hearings in the 1960s, after the Quiz Show scandal, which were “traumatising” for the Nielsen witnesses and put some other ratings companies out of business; Australia’s shift from diaries to people-meters for measuring TV viewing in the early 1990s and then to a different kind of meter and a new operator and owner of the ratings data in the early 2000s.

*Rating the Audience* stresses the frequency of these conflicts and the familiarity of the issues at their heart. Transistor radios allowed mobile listening and made people rather than households the listening units. Video cassette recorders and electronic games players enabled people to do something with TV sets other than watch measured television services. Still, the measurement maelstrom today seems of a different order to those brawls. Some elements of it have been building for decades; others are more recent.

First, what has been building for a long time is the fragmentation of media use. More media options mean smaller numbers tuned in to any one of them. That requires bigger samples if the results are to be statistically reliable and bigger samples cost more money. The internet did not invent this. Until pay TV arrived in the mid 1990s, audiences in the big cities had more radio stations than TV channels to choose from. Radio samples needed to be larger, although the industry’s revenue was smaller. The internet, however, has dramatically increased the electronic choices available to users. Measuring what online and mobile users are doing is more like trying to determine what book they were reading or what record they were playing: no one much bothered to ask that of individual users in analogue media days.

Second, more active audiences strain existing measurement methodologies. New media devices have *decreased* the simultaneous consumption of most media experiences by large numbers of people but greatly *increased* the simultaneous consumption of different media experiences by individuals. So media products like TV dramas now gather their audiences over time – on first release, on catch-up TV, DVD and download, and via apps, all measured in different ways – but particular individuals are members of more than one audience or user...
group at a time – watching a TV program, on Facebook, text messaging.

Service providers, content producers and especially advertisers want to understand not just how many people used their service or watched their show (“consumed their content”) but also what else they were doing at the same time, how engaged they were, if they were letting others know how they felt. For this year’s Olympic Games, the Nine Network is offering its telecast partners and sponsors a “customised real-time cross platform reporting dashboard” integrating “metro and regional TV ratings, online ratings, online video and mobile ratings, App usage, social media buzz and brand health metrics.”

Third, media users engaged in this kind of multi-screen complexity have little time or patience for the kinds of measurement activities used in earlier eras, like completing diaries or pushing buttons on a TV meter. They are tired of being asked, don’t have time to answer and might not think too much before they do. Particularly in some demographic groups, they are less likely to have the telephone landlines whose near universal take-up gave researchers a contactable census from which to draw demographically representative samples. ThinkBox, the marketing body for commercial TV in Britain, recently commissioned a major study that watched its subjects: a research company filmed the living rooms of twenty-three homes and examined over 700 hours of TV viewing for a psycho-psychological analysis of multi-screen behaviour.

Fourth, by leaving their traces with operators like Google, Amazon, Facebook and Apple, new kinds of media use have created immense, proprietary sources of data about user behaviour to rival those collected from structured panels measuring radio, TV and now online use. For these behemoths, panel data offered by third-party providers provides an interesting second opinion about what internet users are doing, not, like TV ratings, a universal currency broadly accepted by everyone trading in their markets.

**SO THE** information age has a paradox: more sophisticated tools enable us to know so much more but increasingly complex behaviour means there is so much more we need to understand. Digital media users are easier to track but harder to follow.

This is not just a battle for the media industry whose users are being measured. Balnaves, O’Regan and Goldsmith stress the public as well as private role that ratings providers have always played. The global financial crisis showed how significant the measures of risk determined by information providers like Moody’s and Standard & Poor’s were to the decisions made by private traders in financial markets, eventually with immense public consequences.

Data about media use will be directly relevant, for example, to evaluations of the National Broadband Network. Just what people choose to do with much faster broadband is plainly a matter of big political significance, given the role the issue played in the 2010 election and the formation of a minority government. Since then, the Convergence Review has recommended that a new category of major “content service enterprises” should become the target for some forms of regulation that previously applied only to TV and radio licensees. One of the elements proposed for determining exactly which “content service enterprises” is the size of their audiences. If the recommendation is adopted, we’ll need an agreed way to measure that.

*Rating the Audience*’s historical perspective on contemporary media measurement highlights both its unprecedented elements and its familiar complexion. By exploring how we got here, the authors remind us of the necessity and the limits of numbers, and of the conflicts and compromises that go into defining those shifty entities sometimes still known as audiences.