The Sternberg Report

PREMIUM EDITION

May 2023 #167

Does Anyone Really Want "Better" Audience Measurement?

By Steve Sternberg

The short answer is no. Sellers want higher reported ratings, while buyers want stability. Neither is conducive to better, more accurate audience measurement. The long answer has a lot of extra words and exposition, but still winds up being no.

Of course, advertisers want the most accurate data possible, but with a couple of notable exceptions, they have traditionally had surprisingly little influence over the way national television audiences are measured.

Nielsen is an industry-mandated monopoly. Over the past 40 years, there have been several failed attempts by various companies to compete with Nielsen. Except for ComScore, none have managed to pose any real threat to Nielsen's dominance. While they have served to pressure Nielsen to make improvements to its own service, the industry (both buyers and sellers) has balked at funding two audience measurement companies that essentially provide the same thing.

Not counting increasing sample sizes, **there have been only two instances, more than 20 years apart, spurred by forces beyond its control, when Nielsen made real improvements to its audience measurement** and how it reported television viewing – in 1987, when the switch was made from the antiquated household meter/diary system to national people meters, and in 2009, when DVRs and time-shifted viewing led to the shift from program ratings to average commercial-minute ratings (C3).

In 1985, when AGB, which had been using people meters in the UK, tried to introduce them into the U.S. market, it represented the first real threat to Nielsen's dominance. Even Nielsen had to admit that this new electronic measurement was far superior to its own meter/diary system, which had considerable human bias – the highest rated networks and programs, as well as shows airing multiple times per week (daytime, early morning, evening news, late night, syndication) benefited when people manually filled out diaries at the end of each week. Recall favors the most popular and longer-duration programming. Facing the possibility of losing clients for the first time, **Nielsen was forced to develop its own people meter**.

Research is traditionally slow to change and adapt to shifting marketplace conditions. Before people meters brought the first major change to TV audience measurement in 30 years, the 3-network, 15-channel, one-screen, live-viewing TV world did not require much change in how viewers were counted. This enabled Nielsen to continue trudging along without needing to innovate too much or pivot too quickly.

In 2007, commercial avoidance through DVR fast-forwarding had become one of the biggest concerns among advertisers. DVR penetration among U.S. TV homes had just hit 20% and was expected to rise dramatically over the next few years. Measuring DVR playback and delayed viewing was now a priority. Virtually every study and survey ever done indicated that anywhere between 70% and 80% of VCR or DVR playback involved fast-forwarding through commercials. But there was no actual data to prove it. Nielsen had never been able to adequately measure VCR viewing and its national people-meter sample had just started including DVR homes. **Nielsen was forced to report DVR playback**, **which led to C3** (average commercial minute ratings up to three days after the original broadcast).

How these two fundamental changes were implemented shows how much the industry's priorities shifted over time – the rapid change in the volume of and access to video content and multi-media devices, as well as delayed viewing options and the resulting changes in consumer behavior, has caused good research to take a back seat to rushing new products to market.

There was a time when any changes to audience measurement needed to undergo extensive vetting before the industry would consider implementing them. **People meters were evaluated by industry committees for nearly three years** – including a full year of side-by-side data to compare it with the old meter/diary method – before agencies and networks agreed to use it as marketplace currency.

Before officially making the switch from program ratings to C3, however, virtually no research or industry vetting was done. This was the first time that marketplace business concerns overrode the necessity for validating research methodology. It also marked the moment that Nielsen shifted away from being a pure research company – at the following year's national client meeting, **Nielsen actually proclaimed that getting a product to market was now more important than validating the research beforehand**.

The major impediment to real improved audience measurement has been that, despite claims to the contrary, **neither buyers nor sellers have traditionally wanted more accurate audience measurement**. Sellers want higher ratings (so they can charge more for advertising) and buyers want stability (so they can project future performance). More accurate audience data is conducive to neither.

Sellers want higher ratings, but more accurate measurement invariably results in declines for the highest-rated vehicles (broadcast networks). You never hear any network complain that Nielsen is overcounting their audience (although for years it did). But every time I can recall reported television usage or network ratings declining suddenly, they blame Nielsen and call for change. There's a reason some networks pushed so hard for viewing in bars and restaurants to be included in currency measurement. Anyone who knows anything about viewing habits or audience measurement knows how absurd it is to think more than a small fraction of people in bars or restaurants watch or hear the commercials. It's not about being more accurate; it's all about higher reported ratings. On the other hand, when Nielsen was unable to report VCR playback, the networks were more than happy to continue using the artificially inflated numbers Nielsen did report (for about 20 years), until DVRs overrode the need to properly measure VCRs.

Any new measurement system that did not provide higher broadcast network ratings than Nielsen, was not able to sustain major media company support. Today, however, with the major media companies owning linear broadcast and cable networks, as well as ad-supported streaming platforms, losing broadcast viewers but gaining more viewers elsewhere may be more palatable.

Buyers want stability, but actual viewing patterns are no longer stable – accuracy is preferred, but stability is required to be able to project future performance (a key element in buying commercial time). Back in the days when there were relatively few viewing options, delayed viewing was not yet a thing, and most content distributors started and ended their seasons at the same time, there actually was a degree of stability to people's viewing habits. If you watched *Seinfeld* on Thursday at 9pm, you tended to do so almost every week. Not so today. Traditional TV seasons, with new series debuting in the Fall and Spring, might still exist for the broadcast networks, but not for other platforms (or for viewers). Streaming services drop new seasons and new shows throughout the year, whenever they are ready, and not always at the same time as the previous season.

Some streamers drop entire seasons at once, some two or three episodes before switching to weekly, and others use the broadcast network once-a-week model. When a new season of *Stranger Things*, or a new show like *Squid Games* or *The Recruit* premieres, my viewing habits immediately change. I'll binge a few episodes a night, and DVR my regular linear shows (which my wife and I might binge two or three weeks later). **My viewing is not nearly as stable or predictable as it was 10 years ago**, and much of my viewing would not be captured by Nielsen's C3 (or C7) measurement.

If you walk into my 23-year-old son's bedroom, you might see him with a muted basketball game on his television set, while streaming something on his laptop, with a soccer game on another screen, and something else on his phone. I have no idea how much of his video viewing would be picked up and accurately measured by Nielsen.

And let's not forget that **media agencies do not want to incur substantial cost increases** for improved ratings systems or multiple currencies that would require significant spending for new or custom databases (and staff) without increasing revenue. Anyone who was in a media agency research group when the industry switched over to buying and selling based on C3, remembers the chaos that ensued during that upfront and for much the next couple of years. There's a reason that what was supposed to be a "one-year band-aid" is still being used 14 years later. Moving to more accurate commercial measurement would require not only significantly more staff, but would produce significantly less stable and projectable ratings.

Before any real improvements to video audience measurement can happen, we need to know the current methodological shortcomings. What can and cannot current rating services accurately measure? While this might sound logical and simple, I can't remember this ever being done (outside of the old telephone coincidental days).

Even when the industry took the major step of moving to commercial minute measurement (C3) in 2009, there was precious little analysis done into the methodology Nielsen uses to calculate these ratings – most people currently at Nielsen don't even know precisely how it's calculated. I remember soon after C3 was implemented, Nielsen put out a note to clients saying that you couldn't simply subtract live viewing from live + 3 (or 7) to calculate the amount of delayed viewing (which most of us and the press were doing). This raised a number of obvious questions. But instead of providing any explanations, Nielsen simply stopped talking about it and never brought the subject up again.

It is actually not that difficult to determine the accuracy of current TV measurement. Fifteen years ago, when I was on the CRE, I proposed a methodology for doing just that. Nielsen didn't want to have anything to do with it, and since they funded our research, it never happened. I've brought this up at least once a year since then.

Here's what I proposed:

- Develop a sample of 100 or so of the top researchers in the industry (buyers, sellers, advertisers), who are adept at detailed work and understand how audience measurement works. Nielsen, ComScore, and others, would meter their TVs, DVRs, laptops, smartphones, and any other video viewing capable devices they own. They would also be given portable meters to measure their exposure to out-of-home video and audio.

- Set up several different viewing scenarios for what and how each participant will view video content and commercials during a single day.

- Participants would then log in their in-home and out-of-home video viewing down to the second (to all content, including commercials).

- Compare their actual viewing to what the meters report. We would then be able to see address weaknesses in audience measurement, how much real-world viewing is not being captured by current methodology, and gauge the differences between such things as exposure (the opportunity to view) and actual viewing, and between individual commercial viewing and C3.

This will tell us what the industry needs to do next. The project could be overseen by a committee of retired research executives or some well-respected company with no current skin in the game.

The problem with this, of course, is that **neither sellers nor buyers have any incentive to actually figure out the shortcomings of video audience measurement**. Sellers certainly don't want to risk finding out how much smaller actual viewership is to their commercials than what Nielsen currently reports. Buyers don't want to have to tell their advertising clients that data they've been using for decades is fundamentally flawed, and they waited until now to do anything about it, when this could have been easily figured out years ago. And both buyers and sellers realize that once you pinpoint specific weaknesses in research methodology, you have to actually fix it – which will be time-consuming and expensive.

There has always been a tendency to start focusing on the next shiny object before actually fixing what is still the biggest piece of the pie – in-home viewing to television sets. But if we don't get that right, measuring real cross-platform viewing will remain out of reach (or just more making up numbers).

It's interesting that even with all the criticism leveled at Nielsen, the company is still used as the base to compare any new potential vendors to gauge the validity their reported ratings. So, if another company reports data that is 30% higher or lower then Nielsen's, it is immediately thought to be "wrong." But we really have no idea if this is actually the case.

It's important to realize that in today's splintered media environment, demographic cohorts are not as likely to have as similar viewing habits as one another as they were in years past. Different companies with different samples will report substantially different results. If the data provided dramatically diverge from one another, none is likely to be seen as a valid alternative to Nielsen as marketplace currency.

A giant set-top sample in the range of tens of millions would be best, but it would still need to be representative of the country at large based on age, income, ethnicity, media device ownership, streaming subscriptions, etc. But even that would significantly under-report viewing sources with a large amount of over-the-air only homes, and would still need a Nielsen-type panel to ascribe demographic data to these homes. But if I had faith that the household sample was actually representative of the country, I would have no problem with developing a panel for demos.

When I was head of Research at major media agencies and negotiated Nielsen contracts, multi-year deals were so much cheaper per year than a one- or two-year contract, that short-term deals were never part of the discussion. I'm sure the same is true for the major media companies, who tend to sign staggered seven-year deals. This, of course, makes it virtually impossible for anyone to come in and actually replace Nielsen. **Until the industry decides to sign one- or two-year Nielsen contracts, we might see some measurement companies provide data used to supplement Nielsen ratings, but a true alternative will remain imaginary**.