

## In Terms of ROI: What Do We Know About Today's TV Ratings Currency? - Bill Harvey - MediaBizBloggers



Published: November 25, 2009 at 05:10 PM GMT  
Last Updated: November 25, 2009 at 05:10 PM GMT

By [Bill Harvey](#)

How accurate are the Nielsen ratings? That question underlies the choices that the industry faces today. Here's what we know.

### 1. CONTAM

A handful of Nielsen fieldmen in the early 60s were caught selling the contact info for Nielsen panelists to TV stations and this precipitated the Harris Committee hearings in Washington. The National Association of Broadcasters and the (at the time) three networks formed The Committee for National TV Audience measurement (CONTAM) to plug the dike with good research before the government stepped in to (innocently/ignorantly) act in the public's interest. It worked.

CONTAM did two very interesting and meaningful studies:

a. Study Number One simply showed that larger samples produced results closer to the results of the largest sample (500,000 diaries) – but with three networks dominating 90% share in 1963, a 2000 sample was terrific and 1000 sample was pretty darn good for measuring the diversity then existing in U.S. TV usage behavior.

b. Study Number Two showed that the people not responding to the Nielsen and Arbitron TV studies tended to be lightest TV users and therefore the >50% cooperation rates existing at the time results in >10% inflation in ratings. (Today the cooperation rates are at less than half that level, so the ratings inflation would probably be greater.)

### 2. Hooper

a. This was a study done in 1968 by a commercially interested party – C.E. Hooper – the inventor of the Nielsen Audimeter – who in the *mid-1930s* licensed that invention exclusively to Nielsen first for radio and then TV.

b. What was done was a telephone coincidental – which at the time was regarded as a truth standard – before the answering machine caused that method to join all other methods as being merely arguably usable for decision making, rather than a truth standard. The reason the coincidental was so highly regarded was that it did not involve

any substantial use of memory but merely asked what people were doing co-incidentally when the phone rang a moment before; and it could be assumed back then (as proven by validation work) that 99.9% of those who did not answer after six rings was probably not home therefore could not be watching TV at home at that moment.

c. What the study found out was that people who responded to a coincidental who were then asked if they would join a meter panel, provided coincidental responses which closely agreed with meter results – but the full coincidental sample findings diverged significantly from those results. In other words, this study proved that non-response bias (i.e. failure to respond, in this case, refusal to accept a meter) was a real effect and that it changed the ratings numbers significantly.

d. The study was done in New York where at the time there were six stations – the big three networks plus three independent stations – New York at the time having therefore the most fragmented audience in television. One of the stations had an 11.5 rating in the coincidental, but an 18.3 rating in the coincidental among those who agreed to join a meter panel. That same station had a 16.1 rating in one of the two N.Y. meter panels and a 14.5 rating in the other N.Y. meter panels (the two N.Y. meter panels were run by Nielsen and Arbitron at the time).

e. This finding along with other findings showed that the willingness to be in a meter panel caused a change in the ratings compared to what was considered truth at the time (the coincidental). The full report as published by ARF available upon request (email me at [bill@traglobal.com](mailto:bill@traglobal.com) if you are interested).

f. In other words, expected truth was 11.5. The ratings services showed the same number to be 14.5 or 16.1, and this inflation was explained by showing that if you asked the diverse sample where the rating was 11.5 to join a panel, those who would join retabulating their coincidental responses had given 18.3. ***The higher number for this station was a function of the preferences of the psychographic group who would join a meter panel.***

g. Among the meter agreeers the coincidental ratings went up to differential degrees for five of the six stations and down for one of them. The ranking of stations was changed considerably by nonresponse bias.

### 3. The Bedrock Finding

a. Gale Metzger in the 1980s put out a thoughtful document which came to be known as The Bedrock Principles of TV audience measurement.

b. This document stated, among other things, that the variation in TV sex/age ratings across programs was 80%+ the result of variation in household ratings, with less than 20% of the variance being accounted for by who happened to be in the room viewing at the time in those homes.

**4. In the late 90s and early 2000s**, TNS in the U.K. said that this percentage was closer to 90% rather than >80%. On the basis of this fact, TNS argued that one ought not tax the industry a fortune for peoplemeter data but instead lower the cost of TV audience research by using set top box data to measure household tuning (accounting for 90% of the sex/age ratings) supplemented by smaller peoplemeter samples to get the sex/age in-room-and-viewing data (accounting for the other 10% variance in the sex/age ratings).

### 5. TNS BBM Arbitron

a. The nonprofit BBM in Canada showed in a landmark study presented at an ARF conference early this century that the first viewer in the room with a TV was pretty well measured by Nielsen-type peoplemeters, but that additional viewers in the room were undercounted.

b. This made sense in that Nielsen annoys viewers in its sample by putting messages up on the screen until the first viewer announces himself/herself by keying in via their button on the peoplemeter keypad. However after the first viewer there is no way for the system to know of the need to prompt again so there are no added prompts and viewers after viewer #1 may lurk with impunity.

c. This past April, a Nielsen study prompted by the MRC verified this finding with a similar Nielsen finding that on average 8% of viewers do not log in and get away with it in the current Nielsen system.

## 6. Today

a. Nielsen shows that people are spending more time than ever with TV at a time when most consumers say that they are spending less time with TV due to Internet usage according to many studies including one done by IBM in 2007. Who knows? Maybe Nielsen is right – maybe not.

b. Set top box data has as many flaws as Nielsen does but at least provides the industry with a second opinion that it has needed for over half a century. That cannot be a bad thing.

c. Set top box data in the hands of the best practitioners can be a better tool than the same data in the hands of others. Is that true? Time will tell.

d. Can set top box data ever be proven *better* than Nielsen? In terms of sample size, yes. In terms of nonresponse bias, yes (very few homes opt out). In terms of response error, yes (passive, no button pushing, outages identifiable and can be taken out of tab). In terms of covering all sets in all homes, not yet, but stay tuned.

e. Does TRA harbor ambitions to displace Nielsen? No. Not our job. Our job is singlesource, and our currency is designed to be used *with* Nielsen. We add the all-important ROI and purchaser targeting dimensions without having to resort to ratio estimation, fusion, or ascription.

f. Should Nielsen use set top data to make its services better? We have suggested that for years and still suggest it today.

*Bill Harvey has spent over 35 years leading the way in the area of media research with special emphasis on the New Media. Bill can be contacted at [bill@traglobal.com](mailto:bill@traglobal.com).*