

Boris Levitan
founder, CEO
boris@immetrica.com
mobile +1 857 891 4000
office +1 617 731 8039
Skype live:immetrica | WhatsApp | Telegram | Viber

Immetrica, Inc.
36 Bellvista Road, unit 16
Brighton, MA 02135
USA

immetrica

more than thirty years of experience
in audience research systems
see the difference™

Crossplatform measurement urges a rethink of intab, reach

Existing definitions are already too diverse for a level playing field. We must adapt them further to work well for video measurement intended to capture all screens

Allience,¹ a joint platform developed by Immetrica and eGlobal,² can measure any screen, anytime, anywhere using smartphones. The smartphone is on or within reach of most people who have one, most hours of the day; it is the closest practical approximation to the often-quoted ideal of one Nielsen client of a measurement device implanted in the sample member.³ Furthermore, a double-digit percentage and increasing share of viewing is done on mobile devices, mostly invisible to conventional measurement technology—but not to Allience, where the measurement smartphone is either also the playback device or close to one.

When we set out to design this system, we confronted one problem that was conceptual rather than technological: the inadequacy of the intab definition. The intab⁴ is the cooperating part of the sample. Most viewing and listening measures are fractions in which the intab is the denominator (as in a rating), or an element of it (as in a share). The required degree of cooperation varies from a few minutes to almost the entire reporting day, but whichever it is, the intab varies directly with the sample size.

Ever fragmenting viewing requires larger intabs to measure. Larger intabs mean larger samples, which cost more money. They must not become unaffordable to clients. That's why than the conventional TAM (television audience measurement) providers' samples are insufficient to measure even the less popular realtime channels, never mind the usually less commonly viewed OTT (over-the-top) programming or online video ads—even if the technical ability to do so exists to any extent.

Smartphone technology, however, depresses the intab just as it is crucial for it to be increased as much as possible. It is normal for a phone to be discharged or powered off for an hour or even more per day. Contemporary lithium ion batteries are rarely generous when new and lose

¹ <http://immetrica.com/Allience.html>

² <http://ecglobalsolutions.com>

³ "Nielsen can count ... on constant criticism", *Variety*, 28 March 2004 <https://variety.com/2004/tv/news/nielsen-can-count-on-constant-criticism-1117902328/>

⁴ "Intab" is the prevalent term in the Americas. In Europe, the term "daily sample" is often used.

capacity quickly as they age. Also some people shut their phones off at work as it is simpler than silencing them.

Some intab standards, especially in the Americas and in continental Europe, require the measurement device to return data around-the-clock or close to it. By these standards, an hour of inactivity in a day would exclude the phone's owner from that day's intab. People-meters have this problem to a lesser extent; they're always on, implying that the measurement is valid even when the users of the TV sets attached to them are out of the home or watching on their individual mobile devices.

But what works for the conventional TAM companies, albeit with an increasing degree of squinting, doesn't work for smartphone-based systems. Even as they must increase their intabs, battery and usage realities depress them.

Furthermore, however acute the need for a rethink of the intab definition is for *average* ratings and shares, it is much more dire for *cume*, also known as *reach*, which affects reach and frequency—a key efficiency measure for advertisers. Here, there was never an international consensus on how to calculate an intab for a multiday reach, and in 2007 Nielsen Media Research in its new NPOWER analytics system gave analysts extremely broad choice—they can now choose any percentage from one day to every day during which a sample member must be intab to be counted in the cume.⁵ This overnight dramatic change destroyed both the inflexibility and the reliability of standards (although we do not rush to say that this is a negative development). Further confusion comes from the fact that Nielsen's online measurement—which started in the earliest days of Web advertising—uses different and less flexible rules.⁶ And aside from the intab, Nielsen has also created an alternative definition of television reach—in which a sample member may be counted more than once.⁷

Again, we do not wish to criticise Nielsen. To the contrary, it should be lauded for offering customers the largest degree of freedom and extensive transparency, much more than is common among currency TAM providers. Nielsen realised early that the new media consumption environment precluded traditional one-size-fits-all straitjackets, as it has in the matter of timeshifted-viewing buckets, and its consistent decisions on these points create a basis for further adaptation by others.

Outside the United States, the reach interval's middle day's intab is used as a starting point, but what are the criteria of the percentage of days in the intab of which the middle-day member

⁵ Nielsen Media Research, National TV Toolbox Online Help, section: Selecting Criteria Using Selection Tabs, page: Unification Page, see Unification: Custom. Some but not all of this freedom of multiday intab definition is also available in Nielsen's Audience Watch analytics system; see Nielsen Audience Watch User Guide, section: About Unified Samples

⁶ Nielsen Media Research, National TV Toolbox Online Help, section: Cross Platform Reporting - TV/Internet

⁷ Nielsen Media Research, National TV Toolbox Online Help, sections: The National TV Toolbox Workspace; Reach & Frequency Schedules and Reports

must be present?⁸ And do the third-party analytics software platforms, which must enforce the data provider's standard, agree with them? And what to do about near-real-time ratings delivery, such as for programmatic advertising markets? How to define an intab then? It cannot be the full-day one, because the day isn't over!

It's a mess. We think we know how to fix these problems and are working on doing so. Obviously we cannot talk about our plans until they're developed, tested and ready.

Here's an example of how the current approaches work for smartphone-based measurement. As a Nielsen NPOWER user could do, we ran nine Coca-Cola spots' reach and frequency, to custom percentages of the on-air days during which each sample member was intab: 0% (the sample member was intab at least one day of the ad's on-air interval, or, in an exception similar to the BARB one, viewed the ad on a day he or she wasn't intab), 25%, 50%, 75% and 100% of the intab days. The last one returned no views. A requirement that the sample member be intab for 25% or more of the ad's on-air days severely decreased the reach percentage.^{9 10}

Why was it so? It's important to distinguish factors that affect any audience measurement system and those that affect a smartphone-based one. In any system, when impressions are filtered out because they were by sample members who were not intab during sufficient days of the ad's on-air interval, so are contributions to the reach. As the required percentage of days rises, the reach drops.

Then there are the novel considerations that we faced with our smartphone-based measurement strategy. Our measurement app is now practically unkillable, but it did not become so until after the last of these ads went off the air. Conquering high battery consumption to a sufficient degree also took until then; because this was a test, the resulting churn of members out of the sample was mostly not compensated with new sample enrollment. Some of the sample continued to evidence those problems even after we fixed them by being slow to install improved versions of the app.¹¹ Both those aspects of our system have markedly improved since we comprehensively rewrote the app in February, after the test described here: the propensity of a sample member to be intab on a given day, and the tendency not to uninstall our app because of its demand on the battery.

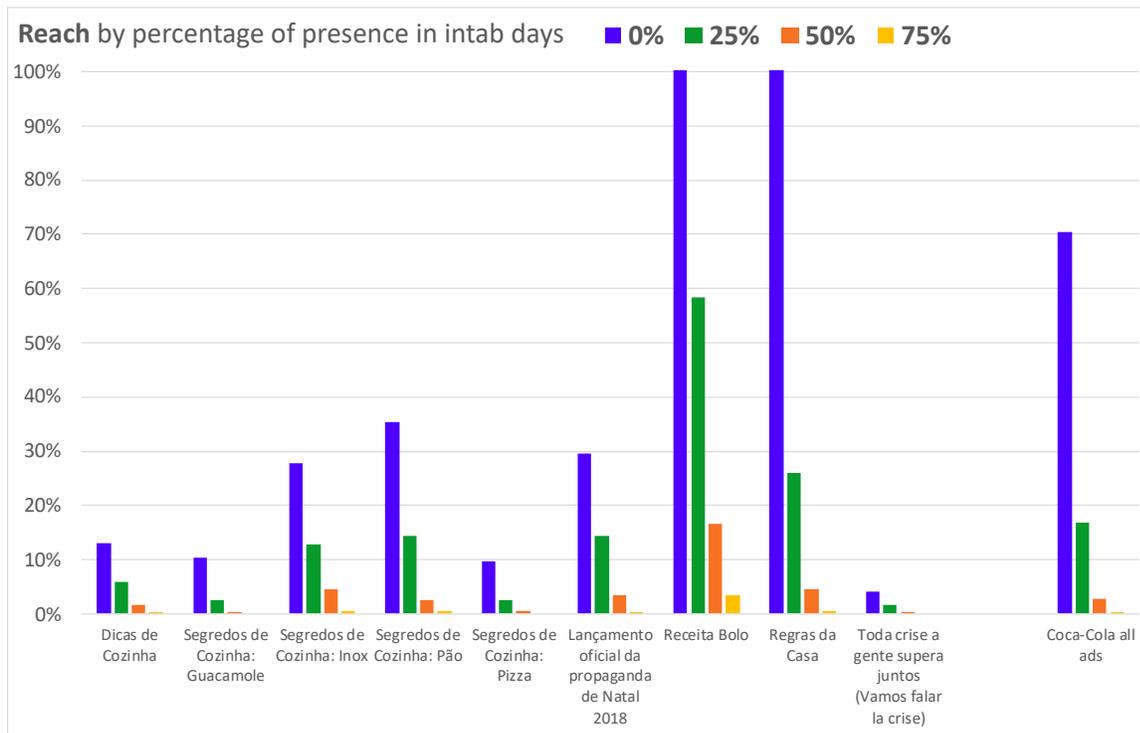
⁸ *BARB (British Audience Research Board) Manual*, §1.12.1, 1.12.3, 1.13. In this implementation, the intab ("daily sample") definition is highly permissive, with the considered members having to be intab only on the middle day regardless of the duration of the reach interval. This decreases the reach rating relative to a less liberal definition. Furthermore, to be intab, at least for reach purposes, a member has to return just three minutes of data for the reporting day—a level of laxity that also inevitably contributes to an undercount. And unlike in Nielsen's current system, there appears to be no escape from those rules.

⁹ Noncooperation for any reason (battery discharged, phone powered off, measurement app crashed) also reduced both reach and frequency.

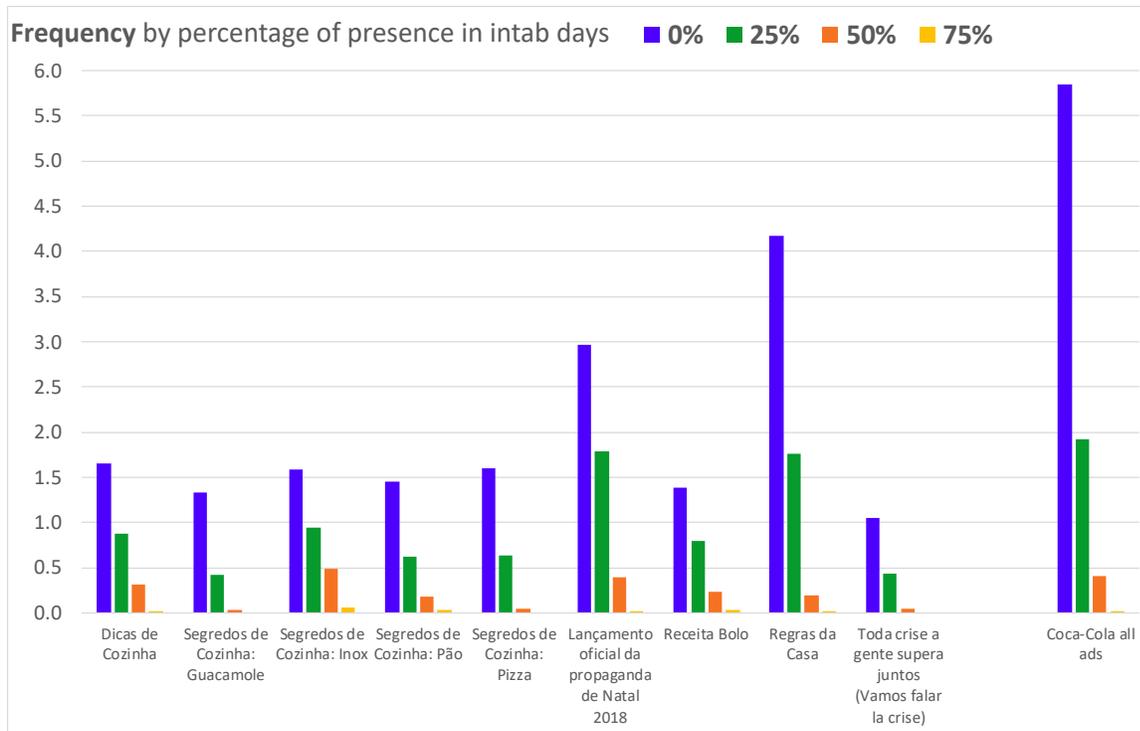
¹⁰ The reach for all nine Coca-Cola ads combined is lower than those of some individual ads because the former has a much longer interval, and thus the a much larger multiday intab.

¹¹ Apparently many smartphone users have automatic app updates switched off, or are running out of drive space so the updates can't happen.

However, the pass/fail definition of the intab remains a problem for any system as capable as ours, and substantial reach undercounts would result from unification to high percentages of multiday intabs, as would the middle-day approach. If we want the comprehensive measurement benefits of mobile-device-based systems, we have to adjust to the concept that periods of noncooperation are normal for them, and algorithms have to be adjusted.



Frequency also depends on conventional expectations. Because it is an average of impressions divided by the sample members in the reach, it is expected to rise with higher cooperation—which tracks with an increased percentage intab during an ad's on-air days. But here we have the opposite effect: the more permissive the reach definition, the higher its average frequency.



We face a choice between conventional, fixed-screen, home-only measurement, where the device has very high cooperation but its user or users are increasingly likely to be away or watching on OTT, computers, mobile devices, or game consoles, and new-era measurement that detects all that viewing but has a perforce lower cooperation rate.