

Selecting the Right Split for an Incrementality Test



By [Yossi Synett](https://kenshoo.com/author/yossi/) (<https://kenshoo.com/author/yossi/>), Vice President, Research @ Kenshoo
August 03, 2018

This is the third post in our series on incrementality testing. In our [first post](https://kenshoo.com/incrementality-testing/) (<https://kenshoo.com/incrementality-testing/>), Omri Cohen introduced some of the reasons why advertisers are using incrementality testing and gave a brief description of what an incrementality test looks like. In our [second post](https://kenshoo.com/measuring-up-5-gaps-of-attribution-measurement/) (<https://kenshoo.com/measuring-up-5-gaps-of-attribution-measurement/>), Tomer Shadi dove deeper into one key motivation for the incrementality testing approach—the failure of attribution to capture incremental value driven by paid media channels. In this post, we’re going to explore incrementality testing methodology—how it differs from more common A/B testing, different approaches for creating test and control groups, and why getting the right test set up is really important.

Let’s start with a refresh on what an incrementality test is from our first post:

“An incrementality test compares the revenue or relevant KPI generated between a test group and a control group. By exposing the test group to an advertising tactic versus the unexposed control group, marketers can easily isolate the affected variables, clearly, assess immediate business impact, and formulate next steps with confidence supported by data.”

From this description we can lay out the **three stages of an incrementality test**:

1. **Preparation** – Split some part of my addressable market into A & B groups

2. **Intervention** – Expose one of the groups to a new variable that may impact performance, allowing enough time for any difference to become apparent
3. **Measurement** – Examine the performance of groups A and B pre- and post-intervention to understand the impact

In this post, we'll focus on the first stage and get you ready to select the best type of split for an incrementality test.

A/B Test vs Incrementality Test: The Crucial Difference

So far, you might be thinking that this all sounds similar to traditional A/B testing (<https://hbr.org/2017/06/a-refresher-on-ab-testing>), where you might test things like subject lines, images, or landing pages to see which variant performs best. But the truth is, an incrementality test is very different. In incrementality testing we are trying to measure the impact of the test on business-level metrics such as revenue, new customers, or site visitors. In traditional A/B tests, it's often about media optimization and we look for the impact on more specific campaign performance metrics such as CTR, attributed conversion rate, etc. As outlined in Tomer's post, we cannot rely on attribution to understand business-level impact and as such **the performance data we use to measure impact in an incrementality test must not rely on attribution.**

Measuring impact on business metrics also means we have to be very thoughtful as to how we set up our split test. In an incrementality test, the split into groups A and B should be done in such a way that the intervention performed on one group will have little to no impact on the other group. Without this guarantee, the results of the test may totally miss or greatly exaggerate the impact of the intervention. In other words, you need a clean split with minimal crossover. Let's examine three of the most common split types used in traditional A/B testing and see which, if any, are appropriate for an incrementality test:

- Auction split
- Audience split
- Geo split

For each type of split we'll evaluate it against the criteria mentioned above as well as more general criteria that are important for all kinds of A/B test splits:

- Ability to measure impact without relying on attribution
- Ability to intervene in one group without impacting the other
- Good correlation between performance metrics in both groups
- The randomness of the split

Auction Split

How it works: An auction-based split randomly assigns a user to group A or B in real time, i.e. when they are about to be exposed to an ad. This approach is used by Google in their Drafts & Experiments product (a recently released beta version provides an option to select a cookie-based audience split as well).

Pros: Theoretically this allows for a totally random split, which is ideal from a statistical point of view, and should lead to a good correlation between the groups.

Cons: An auction-based split has one potential flaw in that the random assignment occurs every auction, so the same user can be exposed to advertising from both groups A & B.

Right for incrementality testing?: NO! This flaw rules out such an approach for any kind of incrementality testing since the probability is high that intervention in one group will have an impact on the other. Furthermore, since there's no clean separation of users between groups A and B, there's really no value in looking at performance data without attribution, as there's no way to associate unattributed conversions or revenue with either groups A or B.

Audience Split

How it works: An audience split assigns users to groups A and B randomly but reproducibly such that the same user will always be assigned to the same group. This is generally done using hashed cookies or other forms of the user identifier.

Pros: Like an auction-based split, this also creates a very random split of two well-correlated groups.

Cons: There are many limitations when it comes to incrementality testing. First, the split is only as good as your testing technology's ability to identify unique users, which is trickier in today's multi-screen, app-filled world. Cookie-based audience splits are likely to assign multiple devices/browsers from the same user to different groups, making true audience-based splits largely possible only for publishers who have a high percentage of cross-device logins (the approach used by Facebook in their lift testing). In order to measure impact without relying on attribution, you need to be able to assign transactions to either group A or B based on the user identifier, without the necessity of a preceding click or impression. Facebook is able to make this assignment for transactions recorded by their pixel but it is not transparent—they don't expose the user level audience assignments to allow third-party technologies to evaluate performance based on data external to Facebook. A further weakness of audience-based splits is that they cannot be used for measuring offline impacts, such as in-store or call center, or offline ads such as TV and radio. This is because it's very difficult to reliably connect online user identifiers to offline transactions.

Right for incrementality testing?: NO! Given the significant amount of drawbacks to this type of testing, we'd have to conclude no on this one, too.

Geo Split

How it works: A geo-based split assigns users to groups using the ability to geo-target both traditional and digital marketing campaigns. Geo splits generally work at the city or DMA level—cities or DMAs are randomly assigned to groups A and B.

Pros: Geo splits significantly simplify measurement since one can easily look at both online and offline transactions by geo without having to perform attribution. Furthermore, since they don't rely on user identifiers, they actually have the potential to reduce the chance of intervention on one group influencing the other. A further advantage is that a geo-based split is highly transparent, meaning you can easily evaluate the results of the test against multiple data sources — even those that weren't considered in the planning of the test. As such, a geo split is the only approach that allows you to measure halo effects such as the impact of investment in one channel on the revenue attributed to another.

Cons: Geo splits are less random than audience or auction-based splits, but using a split methodology that actively tries to create balanced and well-correlated groups overcomes this problem.

Right for incrementality testing?: Yes...if done correctly! At Kenshoo, we've been using the geo-based approach for A/B testing and incrementality testing for more than four years (check out these case studies with [Belk](https://kenshoo.com/pinterest-advertising-belk/) (<https://kenshoo.com/pinterest-advertising-belk/>) and [Experian](https://kenshoo.com/fbaddedvalue2-pr/) (<https://kenshoo.com/fbaddedvalue2-pr/>) to see some of our work). We've applied machine learning approaches to create our own algorithm that creates geo splits with balanced and well-correlated groups. We have found the advantages of the geo-based approach to be significant in our ability to run successful tests that yield meaningful results and stand up to analytical scrutiny.

While this is certainly not an exhaustive list of splits, you won't find another type that lends itself as well to an incrementality test as a geo split does.

Looking Ahead

Keep an eye out for our next post in this series, where we'll focus on different methodologies for the measurement part of an incrementality test, as well as some other things to think about such as: handling conversion latency; approaches to testing multiple marketing tactics, strategies or publishers; and how to use the insights from incrementality testing to affect your business.

Interested in learning more about how Kenshoo can help you better test, execute, and orchestrate your digital marketing efforts? [Contact us today](https://kenshoo.com/contact-us/) (<https://kenshoo.com/contact-us/>) to set up a discussion.

[Request a Demo of Kenshoo \(/contact-us\)](https://kenshoo.com/contact-us/)

Tags: [incrementality](https://kenshoo.com/blog-tag/incrementality/) (<https://kenshoo.com/blog-tag/incrementality/>), [Measure](https://kenshoo.com/blog-tag/measure/) (<https://kenshoo.com/blog-tag/measure/>), [measurement](https://kenshoo.com/blog-tag/measurement/) (<https://kenshoo.com/blog-tag/measurement/>), [Optimize](https://kenshoo.com/blog-tag/optimize/) (<https://kenshoo.com/blog-tag/optimize/>), [xTLx](https://kenshoo.com/blog-tag/xtlx/) (<https://kenshoo.com/blog-tag/xtlx/>).

< [Previous Post](#)

[\(https://kenshoo.com/5-signs-youre-underutilizing-bing-ads/\)](https://kenshoo.com/5-signs-youre-underutilizing-bing-ads/).

[Next Post](#)

[\(https://kenshoo.com/making-good-impression-socially/\)](https://kenshoo.com/making-good-impression-socially/) >

Subscribe to Updates

Submit →