Test What Matters:

Level-Up Your Product Experiments with Behavioral Data

Amplitude

In an era of rapid innovation and deployment, digital-first companies must develop a product strategy that involves effective A/B testing and experimentation. If the turnaround time from “hypothesis to change” is too long, businesses risk losing both customers and revenue.

To reduce these risks, product teams run A/B tests to drive better decisions. However, too often, these tests don’t have a meaningful impact. Existing experimentation solutions don’t provide a clear way to identify the user actions within the customer journey. This leads to misguided attempts to improve the customer experience, such as endlessly testing copy changes on a single screen.

The goal of A/B testing should be to influence user behavior in such a way that it drives your desired outcome. The best way to do this is to examine the behavior of your existing customers that have already achieved the end result you seek. From there, you can design experiments to nudge other users to perform the same actions within your product. It’s a far more impactful strategy than the small tweaks that most teams are used to.

By leveraging behavioral data, you can level up your experiments through better hypotheses, better targeting, and better performance analysis. The result is more meaningful outcomes and the ability to conduct experiments that ultimately determine the best experience for your customers.

Today's Challenges with Experimentation

If behavioral data never drives experimentation, you will be left with limited, surface-level insights about your customers — reducing your ability to target them. Many product teams today struggle with a lack of integration between A/B testing and behavioral data, which negatively impacts the quality and scope of the experiences teams can test and configure.

Your Hypotheses are Based on Limited Data

Without behavioral data, it's nearly impossible to form strong hypotheses around what will drive users to take desired, positive actions. Teams often have to rely on data from outside sources, such as third-party data or web analytics. While this data may contain purchase information or traffic, it lacks the context of behavior.

Take, for example, customer experiences that occur across multiple devices. Signup or subscription behavior may begin anonymously, so the customers end up with multiple profiles. Product teams believe that they are examining the behavior of two different users when in actuality, it is a single user. This leaves gaps in analyzing the results of the experiment and makes it challenging to see the whole picture.

To further complicate matters, companies will face even more limitations as both Google and Apple have announced recent privacy changes. Teams that relied on cookies or tracking IDs for experimentation will need to rethink their approaches.

Only 1 in 8 experiments produce significantly positive results.

source

39% of digital businesses say they struggle to optimize their offerings because they are unable to analyze the full customer experience (e.g., across devices, across products).

Source: **Harvard Business Review Analytic Services Survey, September 2020**

Without access to first-party behavioral data, teams are limited in their ability to answer the question: "What will improve the user experience?" You'll have little to no data for forming a hypothesis around what will solve user pain if you don't have analytics around the complete user journey.

You Can Only Make Small Bets

Every experiment you conduct is a bet that you are placing on its outcome.

"(Bets) help you learn about the things that matter," says Fareed Mosavat, former director of product at Slack. "Bad experiments... adjust things around the edges in an attempt to improve performance in a marginal way. They steal time, energy, and resources from validating more meaningful bets."

Too often, teams focus on small bets, like tweaks to copy or turning a feature on-off. While small bets may manage risk within a release, they rarely result in big payoffs.

Teams may also run small tests in low-traffic areas because they cannot segment their experiments to groups of users based on behavior. With this approach, it may take months to see statistically significant results.

Instead, teams should focus on large bets that they can test quickly and efficiently. But product teams often lack the engineering resources needed to take this testing approach. The tools and processes may lack an end-to-end workflow for experimentation or have a gap in designing consistent tests and performing analysis.

In the end, many teams are stuck. Small bets won't make a dent, and executing large bets seems like a mountain that is impossible to climb.

You Cannot Target Relevant User Groups

Pushing out experiments to random groups limits what you can glean from the results. You may encounter users that would never engage in the behavior that you are testing. Or you unknowingly put the majority of your target users in your "control" group, which means they never see the change.

Likewise, you'll limit user insights if you can't link the same users across multiple devices. This limitation in data poses a potential risk that the same user has one experience on a mobile device and a different experience on the web. Your results instantly become less accurate because you have no way of knowing if users saw one or both of the experiences.

For meaningful experimentation, you must measure results through the lens of your users' qualities and product usage. LinkedIn assesses the impact of testing for different groups, such as users who are all from the same country or use the same social network.

Push tests out to users who match your target profile and are likely to interact with your proposed product change, whether it's improved navigation, a new CTA, or a product recommendation. Targeting based on behavior means that you can measure the experiment across segments to see if a meaningful difference emerges for key customer groups.

Results Lack Clear Direction for Next Steps

A good test yields a clear outcome of the experiment and points you toward the next course of action. Unfortunately, most experimentation tools only offer primary analysis.

Teams can identify the basics, such as “Which variation is winning and by how much?” But primary analysis can only show teams results within the bounds of the test, making it difficult to understand if there are downstream effects on engagement or conversion. A feature designed to increase engagement may lead to short-term improvements for customers but then has an unintended consequence of lowering long-term retention.

Primary analysis may show that metrics have improved, but without secondary analysis, this approach is too focused on short-term gains rather than solving an underlying problem.

Secondary analysis digs deeper into the broader implications of the experiment—such as the test results’ implications for retention, funnels, and calculated metrics.

Secondary analysis enables teams to go beyond questions like “Which is the winning variant?” to actually answering “What is the best experience for this customer segment?”.

Primary Analysis:
Answers basic questions about an experiment

Secondary Analysis:
Goes deeper into retention, funnels, and calculated metrics

Think Bigger: Frameworks for Experimentation

Your product experiments should be designed to resolve the pain points of users who don't reach the actions that you are expecting them to take. By comparing the behavior of different segments of customers, you can uncover new opportunities to guide users to value and pull more precise insights about how your product resonates with them.

At the forefront of any framework should be the mantra, "Build. Test. Learn." Outline the scope of what you are hoping to gain, develop experiments around those goals, and use the results to inform your next steps.

Revenue: Uncover New Sources of Monetization

By digging into the behaviors of power users, you can better understand the activities that lead to revenue.

Keep in mind, some of these actions might be less obvious or may not have an immediate impact on revenue. Power users might engage in a lot of social sharing or content creation before they significantly contribute to revenue. In other cases, power users themselves may be outliers in terms of transaction value or frequency.

You may also uncover customers that don't fall into your typical power user profile yet still provide a consistent source of revenue. By defining multiple power user personas, you can broaden your experimentation strategies.

A framework centered around revenue would use behavioral data to unlock ways to further monetize different segments of users. Your tests may include:

- Researching discounts with power users to determine price sensitivity
- Rolling out new features to high-value subscription customers that are good candidates for additional revenue streams
- Encouraging small spenders to form consistent habits as part of a growth strategy

Behavioral data tells you which users take certain actions that lead to revenue generation. Your experiments can nudge other users to the same behavior or make additional bets with your consistent revenue-generating customers.

Value: Improving the User Experience

To maximize the value your product is delivering to your customers, you have to understand why and what behaviors are associated with customers who stick to your product and with those who drop off.

A great tool to do this in Amplitude is called Journeys. Journeys helps you examine how users behave at different points in the funnel to see what drives conversions and what's getting users lost.

The framework, designed to improve the **value** your users see in your product, creates behavioral experiments to see what it takes for users to reach their "aha moment."

- Explore different experiences based on how frequently users engage with your product.
- Test different onboarding paths based on campaigns or channels.
- Educate users on more sophisticated functionality in the context of the features they have already adopted.

Your best customers will have certain characteristics in common, and uncovering these high-value behavioral traits can lead to experiments that encourage other users to do the same.

Growth: Reaching New Audiences

Your growth strategy isn't limited to those who are deeply committed to your product or pay a lot of money. Champion users can also be a significant source of new growth for your product by driving other users to engage more deeply.

According to a survey conducted by Twitter and marketing data analytics leader Annalect, 49% of consumers depend on influencer recommendations. You can find these "influencers" within your own user base by examining their product behavior.

When you focus on **growth**, you want to hone in on the experiences within your product that drive overall adoption and loyalty. Champions of your product have powerful messages to share that may capture the attention of new potential customers.

- Research your product evangelists to understand what makes their experience great.
- Make it easier for these users to invite new users or create a positive review.
- Test new experiences among these users since they are more likely to spread the word.

By testing their potential audience reach and ability to share, you can measure the results of additional growth or new signups.

8 Core Principles for Your Experimentation Strategy

With a framework in mind, you can begin to think about how to move from small bets to large ones. Use these core principles to guide experimentation. By addressing each one, you can create a playbook for designing tests that achieve meaningful results with clear next steps.

1 Hypothesis

Start with a hypothesis that is based on behavioral data. Look at the activity of your customers who are converting, and form a hypothesis around what would nudge more customers to the same behavior.

Example: Research user behavior to create a hypothesis about a loyalty program model that increases engagement.

2 Variation

Understand the normal variation within a metric that you want to test and how it needs to move before you ship a feature. Avoid testing a metric with high variation because it will be hard to tell whether your test is causing shifts in the metric.

Example: Remote change of a primary CTA for new users versus subscribers. Look at the variations in how segments of users respond to a particular CTA before modifying it.

3 Identity Resolution

Ensure your users can be retroactively merged across all channels as you conduct your experiment to ensure consistency and accurate results.

Example: Conduct signup and subscription discount experiments and be confident that users are seeing the same price whether they are on mobile or web.

4 Context

Run experiments across products and devices to see if they require special nuances or limit to a certain segment, to understand how audience needs vary.

Example: Channel-based form optimization may have form fields adjusted based on UTM parameters.

5 Iteration

Don't assume your experiment is complete once you've hit statistical significance. Figure out what you learned and determine what you can iterate on to test next.

Example: Alleviate distractions from the signup funnel by removing or reordering the different elements.

6 Segmentation

Group users by behavior (for example new users vs. power users) to give you a better understanding of that experiment's impact on the behavior that you are trying to influence.

Example: Test improved navigation for power users or simplified navigation for new users.

7 Predictions

When you group users into cohorts based on their past behavior, you can begin to form predictions about their future behavior. Predictions allow you to target the users most likely to engage in the behavior you are testing.

Example: Use predictions to test changes in the UI based on users' likelihood to use a feature or make a purchase.

8 Recommendations

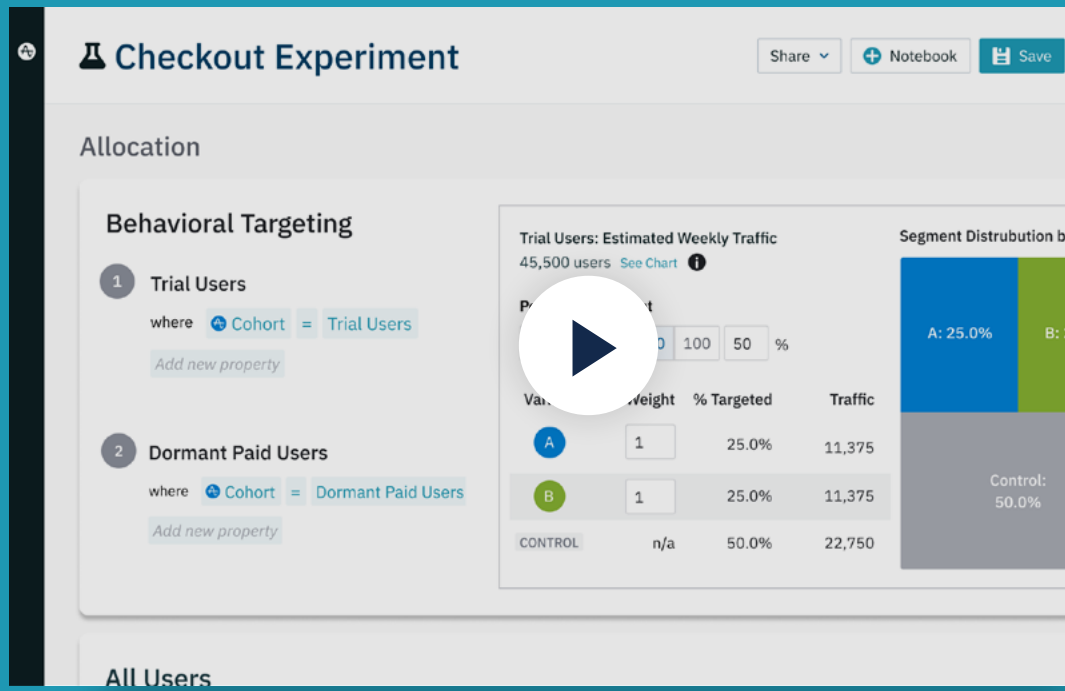
You could take a heuristic approach where users independently discover your product's value. But then, why bother with experimentation? Use your behavioral insights to make recommendations to users that keep them engaged in your product and increase value.

Example: Test add-ons, cross-sells, upsells, and improved messaging based on past behavior or the behavior of other users with similar behavioral patterns.

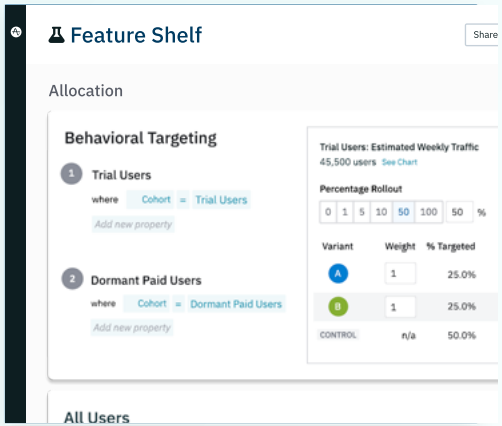
Challenge Accepted:
Introducing

Amplitude Experiment

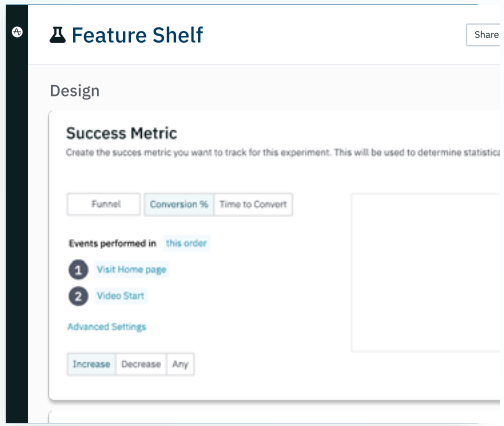
Amplitude Experiment is the first experimentation solution powered by analytics and customer behavior.



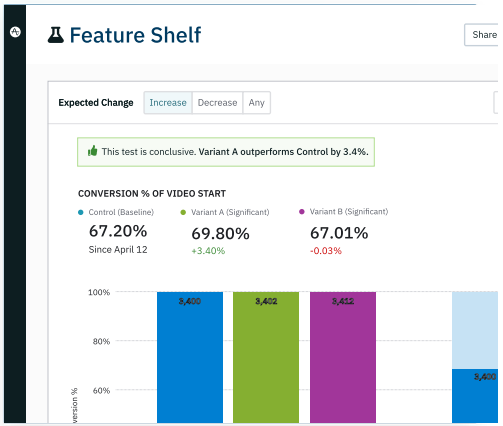
Amplitude Experiment’s end-to-end workflow integrates customer data into every step from generating a hypothesis to targeting users to measuring results.



Design better tests by starting from better hypotheses.



Rollout to custom segments and platforms without extra engineering work.



Learn the true outcome of experiments by seeing the downstream impact.

Find out how Amplitude Experiment will level-up your testing.

info.amplitude.com/experiment-demo

AmplitudeAnalytics amplitude-analytics amplitude_HQ