

EBOOK

Top 10 takeaways from running 127,000 experiments

Contents

p3	Introduction
p5	Test impact: Focus on learning rates, not win rates
p6	The velocity conundrum: Speed is nice. Impact is nicer
p7	Rethinking metrics: Identifying what drives sales and revenue
p9	The uplift decay challenge: Focusing on long-term impact
p10	Complexity is your friend: The power of multi-variant testing
p11	Highest returns: Digital commerce revenue says Hi
p12	Personalization: Turn visitors into customers with 41% more impact
p14	Developer resources: Running more tests doesn't always lead to uplifts
p15	The power of understanding your data: Sharper analytics. Better insights.
p17	Centralized or decentralized: Ways to structure your experimentation team
p18	Conclusion

Introduction

There is still limited public information available about what makes tests successful and how to grow your program. You need resources to prove experimentation value. But without proving value, you don't get resources. It is hard to know where to start.

That's why we've got you covered. We crunched the numbers - from 127,000 tests in fact - to bring you insights that'll help you scale your experimentation program through strategic experimentation choices.

Whether you're just starting or looking to scale your program, these 10 data-backed takeaways will guide you toward more impactful, revenue-driving experiments.

The takeaways are here. But from where?

- Data from 1.1k companies
 - Analysis of over 127K experiments
 - Excerpted Optimizely analysis
 - Customer interviews, case studies, and surveys
- 

Top 10 takeaways

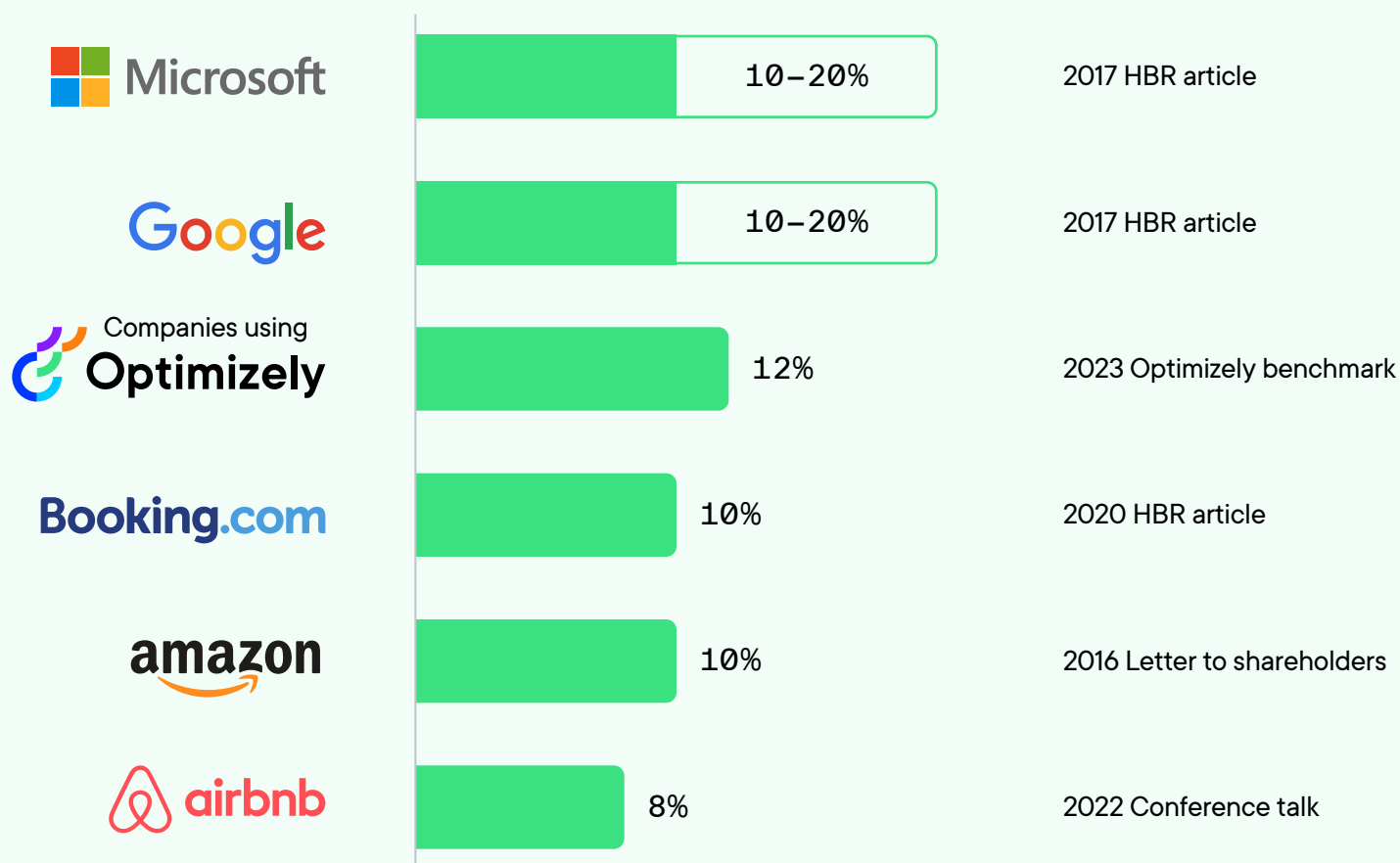


01

Test impact: Focus on learning rates, not win rates

Around 1 in 8 experiments will win on the primary metric

Experiments achieving a statistically significant improvement, self-reported



STAT

Only 12% of experiments win on the primary metric.

Takeaway

There is nothing inherently better about a “winning” test than a “losing” test because both teach you something and both ultimately result in revenue benefits in myriad different ways. Focus on learning rates, not win rates.

You need to start framing the value of experimentation in terms of uplifts and translating win rates into expected impact per test.

Real-world example

[Hunter Engineering](#) boosted leads by 88%, traffic by 72%, and monthly order volume by 20%.



TIP

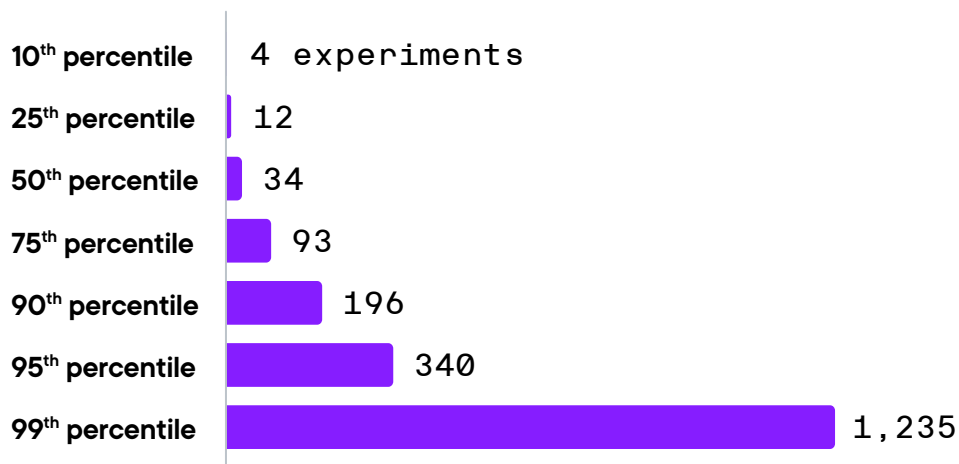
Let customer engagement drive your decisions, stop guessing what will convert.

02

The velocity conundrum: Speed is nice. Impact is nicer.

Experiment velocity by company

Experiments created in 2022, over 900 companies



STAT

The median company runs 34 experiments per year. The top 3% of companies run over 500.

Takeaway

As you scale, shift focus from velocity to impact. Balance quantity and quality. Prioritize high-impact experiments over simple, high-volume tests.

You should:

- Have sufficient developer resources with a diverse portfolio of iterative changes (pricing, discounts, checkout flow, data collection, etc.)
- Document every change and improve how users interact with your website/app.
- Choose experimentation metrics that impact program quality and adoption.

Real-world example

Thinking about brushing up on your conversions? [Quip](#) polished their A/B testing speed by 40x and saw a 4.7% boost in order conversions using Optimizely Web.



7

TIP

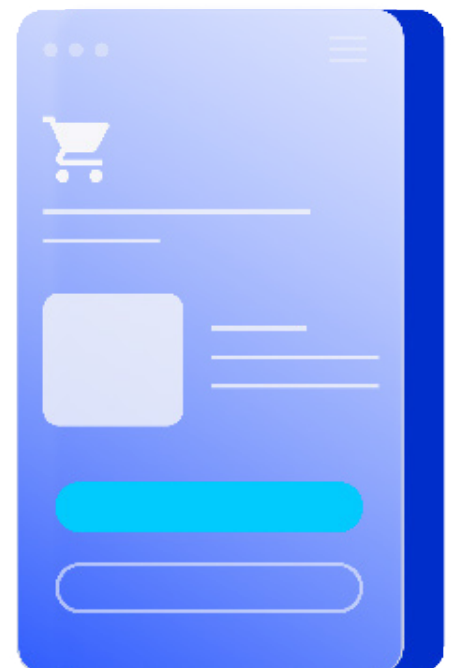
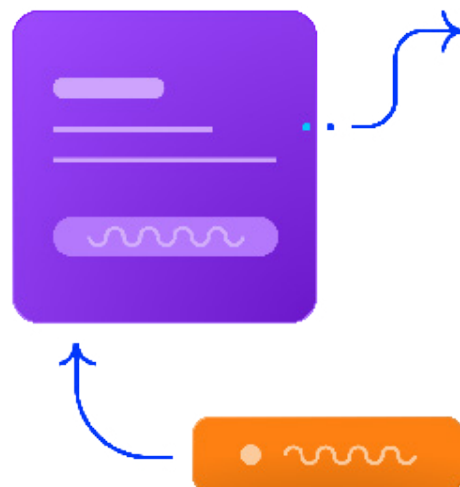
A multi-armed bandit model can help you attain the fastest, most statistically sound results while targeting precise audience segments.

03

Rethinking metrics: Identifying what drives sales and revenue

STAT

90% of experiments target five common metrics, but three have a low expected impact.



Primary metric share and expected impact across all experiments

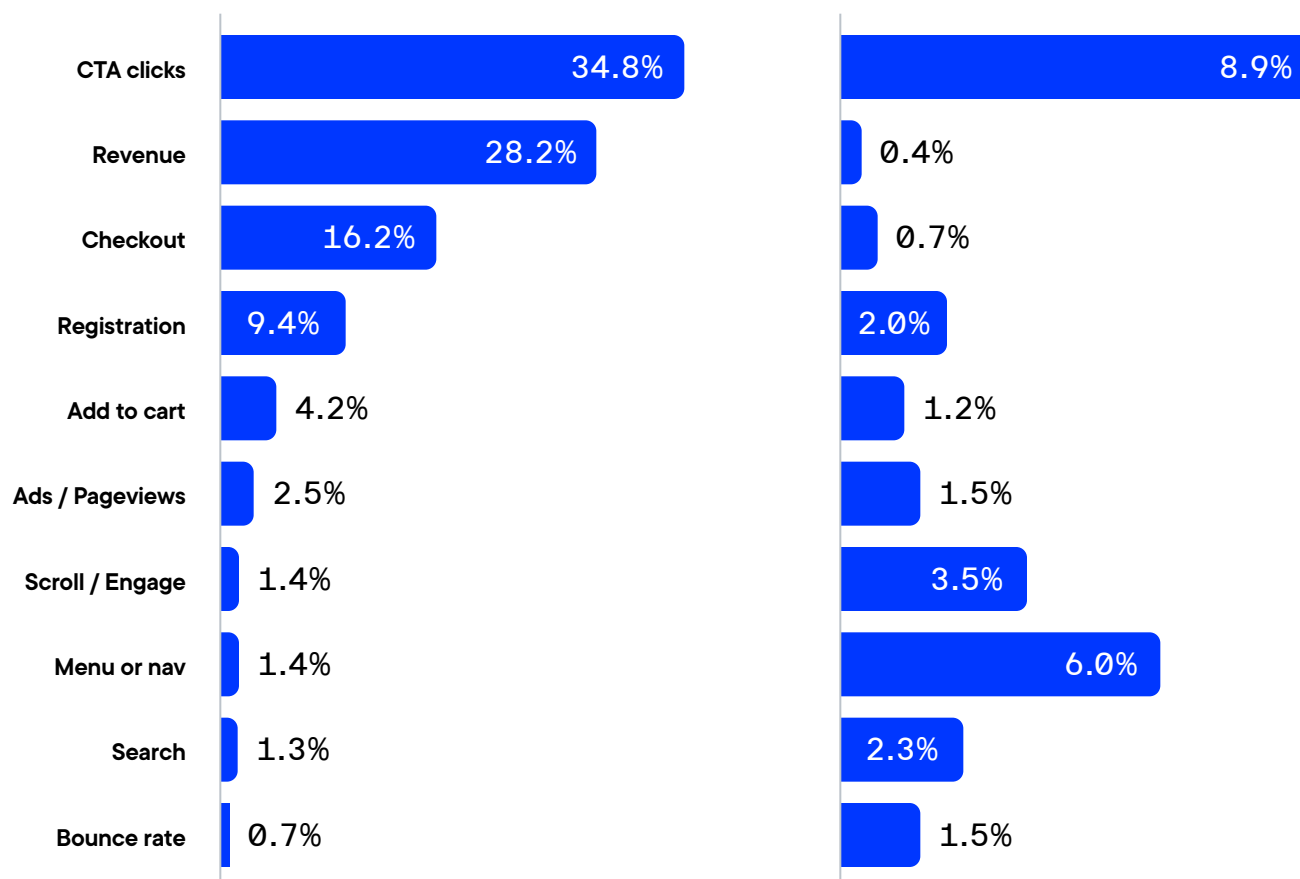
True experiments, n = 1.1k companies, n = 40k experiments

Primary metric share

(% of all experiments)

Expected impact

(Win rate x uplift)



Takeaway

Standard metrics might be limiting your potential gains. Focus on metrics that affect each decision point and lead visitors to the buying moment.

Better metrics → Bigger uplifts → Fatter bottom line and more sales

Real-world example

Just ask [Wonderbly](#), who received 18.1% more orders by tracking the right metrics.



TIP

Have one configurable dashboard that allows you to make statistically sound decisions as soon as the data starts coming in.

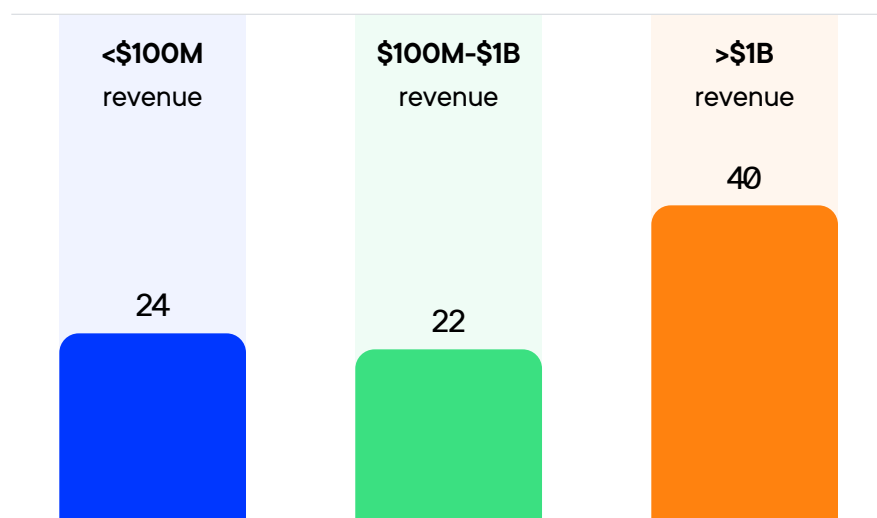
04

The uplift decay challenge: Focusing on long-term impact

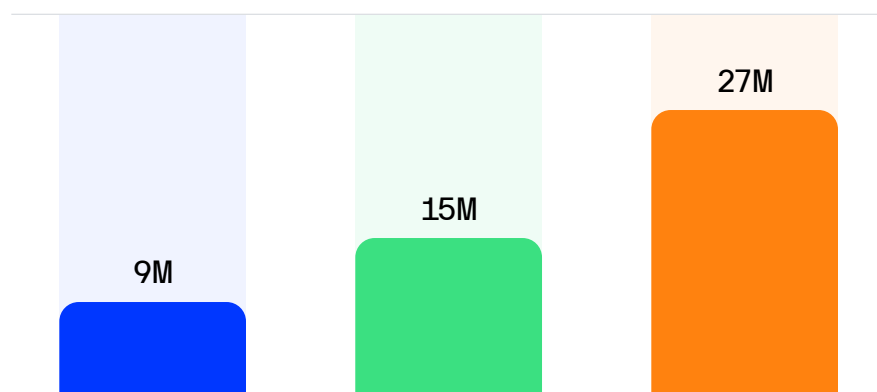
Company revenues, traffic and program metrics
Experiments in 2022, Similarweb traffic data for 2022



Annual launched experiments



Average monthly sessions



STAT

Most experiment uplifts decay to 80% of their first-month value after a year, except for revenue, which retains 91% after a year.

Takeaway

The impact of your experiments may diminish over time. Factor uplift decay into long-term planning. Regularly reassess and refresh experiments to maintain impact.

Real-world example

[Prezzee](#) improved sign-up rates by 580%.

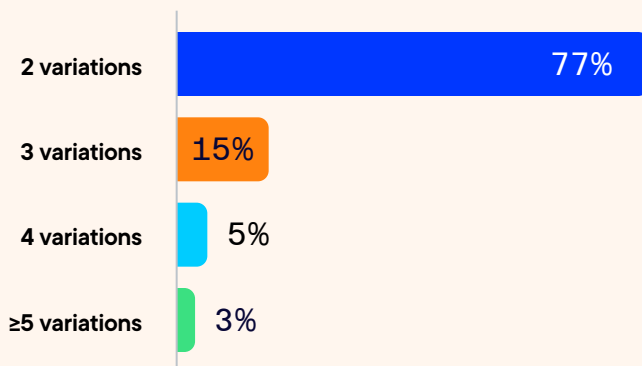
05

Complexity is your friend: The power of multi-variant testing

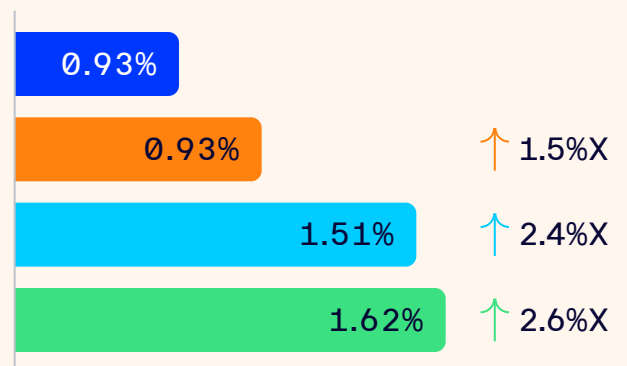
Experiments by variants versus expected impact on checkout

n = 127K tests, variations includes original (2 Variations = Original + 1 Treatment)

Share of all experiments



Expected impact on checkout



STAT

Less than 10% of experiments test 4 or more variants, yet those experiments are more than twice as impactful compared to A/B.



TIP

An effective visual editor helps you target any element on your page and preview changes on the go so you can make multiple changes whenever you want.

Takeaway

The highest uplift experiments worldwide have two characteristics in common: they test a higher number of variations and implement more complex changes.

Simple A/B tests lead to leaving money on the table. Don't shy away from complexity – multi-variant tests often yield higher impact results that'll make your CFO smile.

Real-world example

[Creative Co-Op](#) increases transactions by 19%.



06

Highest returns: Digital commerce revenue says Hi

Takeaway

Although revenue is the most valuable business metric, businesses stand to experience greater test impact by focusing experiments on improving micro-conversions, such as getting more users to search, add to cart, and register accounts.

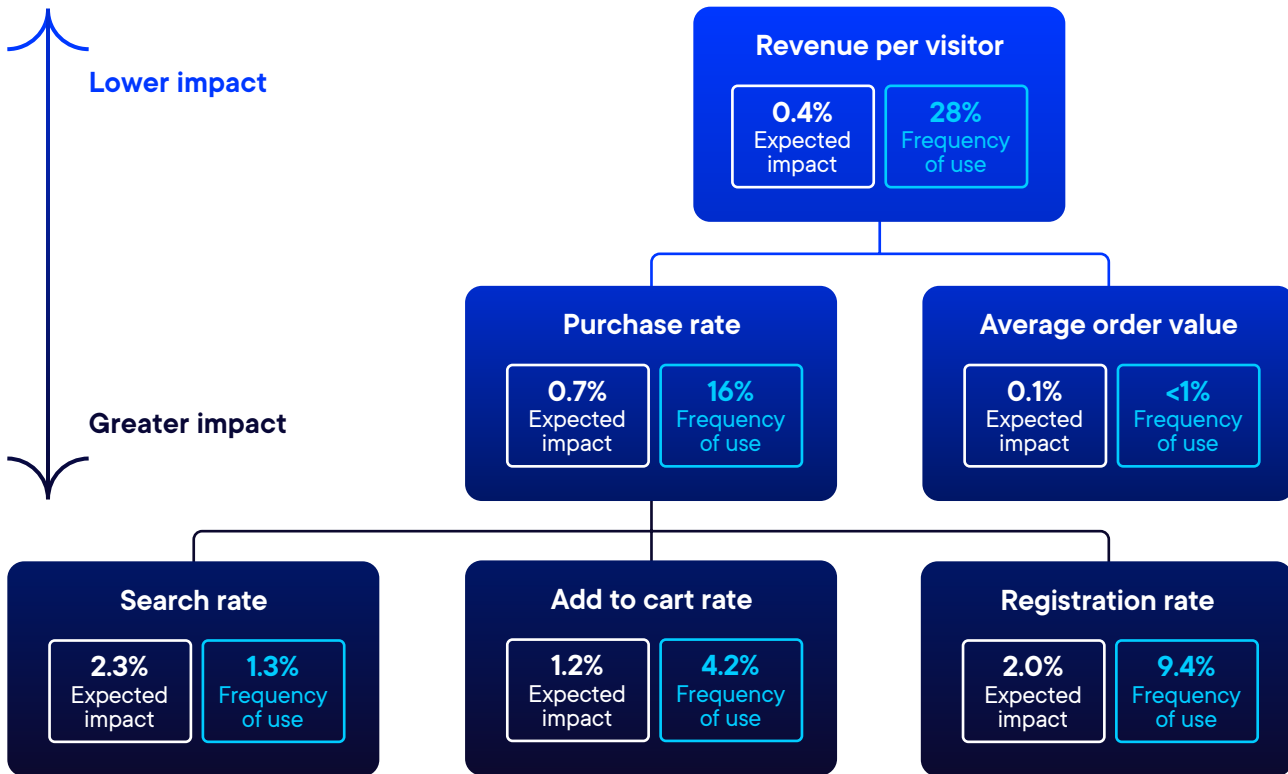
Real-world example

[Tapestry](#) demonstrated a tenfold increase in the scale of their testing program.



TIP

Optimizely allows you to experiment via A/B or multi-variant testing on any channel or device with an internet connection.

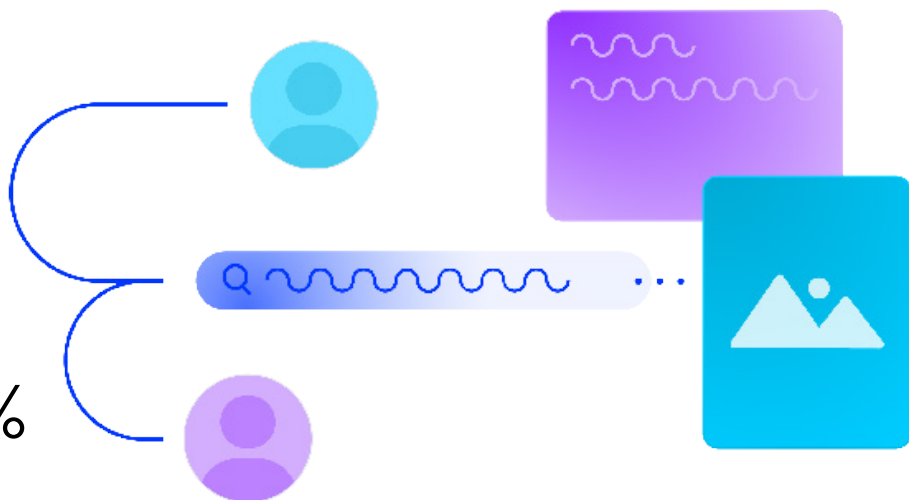


STAT

Digital commerce has the highest returns from experiments that target goals which are early in the shopping funnel.

07

Personalization:
Turn visitors into customers with 41% more impact



Win rates by whether experiments use audience targeting

Win rates on primary metric for true experiments, n = 127k experiments



Average uplift by whether experiments use audience targeting

Average uplift on primary metric for true experiments, n = 127k experiments



Expected impact by whether experiments use audience targeting

Expected impact on primary metric for true experiments, n = 127k experiments



STAT

Personalized experiments generate a 41% higher impact compared to general experiences.

Takeaway

One size fit all is no longer a viable digital marketing approach. You can't just push the same website recommendations to a broad audience.

Tailored experiences aren't just nice to have – they're rocket fuel for engagement and conversion rates. Make personalization a cornerstone of your experimentation strategy.

Real-world example

[Calendly](#) scheduled a date with success, booking a 5% increase in conversions and saving 312 hours of development time annually with Optimizely's experimentation platform.



TIP

Optimizely combines your customer data, powerful testing, and AI-rich personalization capabilities to help marketers create 1:1 experiences that drive revenue and brand loyalty.

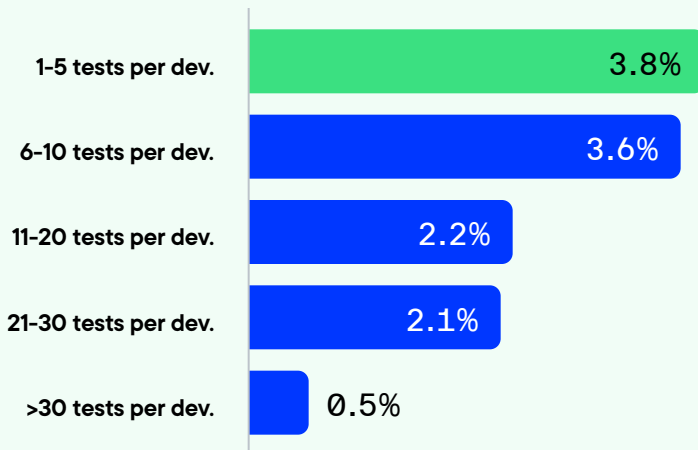
08

Developer resources: Running more tests doesn't always lead to uplifts

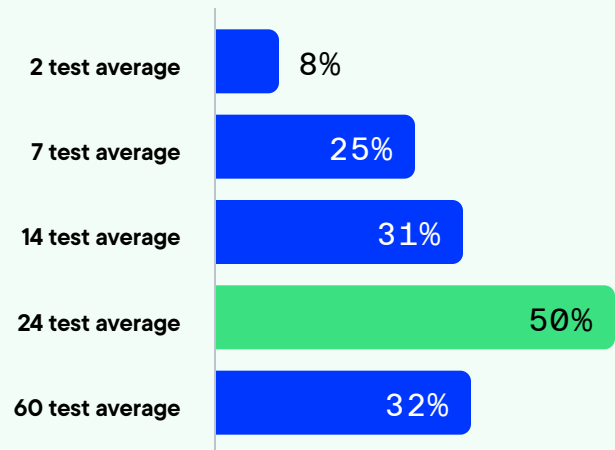
Expected impact based on annual experiments per developer

True experiments run in 2022, n = 392 companies with 12+ experiments in 2022

Expected impact / test



Total expected impact



STAT

The highest expected impact occurs at 1-10 annual tests per engineer. Beyond 30, impact drops by 87%.



TIP

No code changes are required for Optimizely Web Experimentation. You get project-level configuration options that allow you to choose which platform you want to use.

Takeaway

Overloading your development team can decrease experiment effectiveness. Focus on fewer, high-impact tests rather than numerous small changes.

Volume at the cost of quality can harm performance and the expected impact of your experiments. Set clear hypotheses for your experimentation team and provide them autonomy for decision-making.

Real-world example

More than 20 teams at [DocuSign](#), including marketing, product, design, and engineering, actively utilize Optimizely for experimentation.



09

The power of understanding your data: Sharper analytics. Better insights.

Takeaway

Analytics isn't just an interrogation of data. It's about having the ability to think critically through assisted use of data.

Experimentation can help you:

- Break down business problems into parts and steps
- Design analyses and experiments
- Recognize and avoid cognitive bias

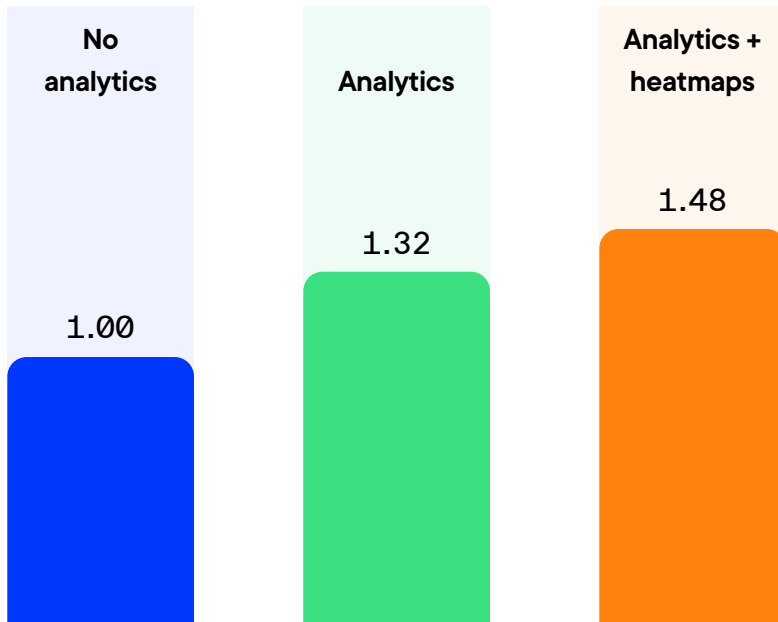
Invest in robust analytics tools to dramatically improve your experimentation program.

Real-world example

Aura increases the average order value by 15%.

Experiment performance with or without analytics in the tech stack

Experiment win rate indexed against “No analytics”, Optimizely +
Built with data, 2018 n = >1,000 companies



STAT



Teams with analytics outperform those without by 32% per test. Adding heatmapping increases success by an additional 16%.

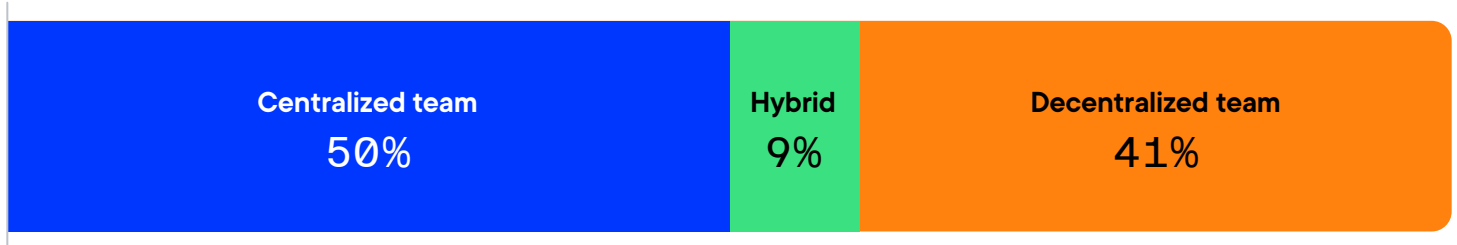


TIP

Our tools count conversions based on unique visitors to your page through our advanced analytics system, meaning users are only counted once regardless of how many times they trigger conversions.

10

How to structure your experimentation team



Takeaway

Most companies begin with a small group of experts of a centralized CRO team. Over time, these experts branch out and create more testing teams throughout the business.

At larger scales, companies are evenly split between centralized, decentralized, and hybrid models. Company approaches are based on the talent, business needs, and culture of the organization.

STAT



Large programs are evenly split between centralized and decentralized teams, with limited performance differences observed.



TIP

Implementation is as simple as pasting a snippet of code on pages where you would like to test or track goals. With Optimizely, run tests as quickly as they are added.

Conclusion

Experimentation is about driving meaningful business impact, not just running tests. By applying these learnings and insights, you can scale your program into a strategic driver of growth and innovation.

Remember, the goal is to run smarter, more impactful experiments that directly contribute to your bottom line. Start by implementing one or two of these takeaways and gradually build them into your experimentation culture.

Ready to take a deeper dive into the world of experimentation?

Here's what we recommend:



- Full report: Evolution of Experimentation
- How to start running perfect experiments
- How to build a winning experimentation program
- Product experimentation guide
- AI in product development

And if you're ready, Optimizely's comprehensive platform offers all the tools you need to implement these best practices, from advanced analytics to AI-powered insights.



Schedule a demo at optimizely.com/get-started today to unlock the full potential of your experimentation program.

At Optimizely, we're on a mission to help people unlock their digital potential. With our leading digital experience platform (DXP), we equip teams with the tools and insights they need to create and optimize in new and novel ways. Now, companies can operate with data-driven confidence to create hyper-personalized experiences. Building sophisticated solutions has never been simpler. Optimizely's 900+ partners and 1100+ employees in offices around the globe are proud to help more than 9,000 brands, including Toyota, Santander, eBay, KLM and Mazda, enrich their customer lifetime value, increase revenue and grow their brands. Learn more at **[Optimizely.com](https://www.optimizely.com)**