Launch and monitor your experiment

Skip Ahead

Keep an eye on health metrics
Pause and re-launch a running experiment

This article will help you:

• Check a running experiment to ensure that reporting is reasonable
• Identify if you’re delivering a bad experience to valuable visitors, and take action
• Pause and re-launch a broken experiment

Once you launch a test, it’s tempting to keep checking its progress because you want to implement a winning variation as soon as you can to improve your business or stop an experiment that’s losing to test a new hypothesis.

Our proprietary Stats Engine lets you to safely peek in on running tests in Optimizely whenever you want. Unlike in classical statistics, where monitoring your test increases the chance of calling a false positive and declaring a winner where none exists, results in Stats Engine are always valid. Peeking can actually help you monitor the health of your experiments.

But it’s still not efficient to constantly check your results. Testing should be treated as a standardized, scalable process; not a project you’re always monitoring. How often should you check your test? What metrics should you look for and how do you take action based on what you see?

We recommend that you check each test after launch at least once, to make sure reporting looks reasonable and doesn’t point to any technical issues with goals or targeting. Once you’ve determined that your test is running properly, feel free to let it run until the projected time-to-completion you predicted in your experiment plan.

What you need to get started:

Materials to prepare

• Individual experiment plan
Approval from your QA team

People and resources

- Program manager, responsible for updating the team and monitoring data
- Executive sponsorship

Actions you’ll perform

- Share the experiment plan with QA, marketing, support, and development stakeholders
- Launch the experiment
- Monitor results
  - Optimizely results, with a focus on the significance and lift parameters that you set in your test plan
  - Evaluate performance for different audience segments
  - Third-party integrations

Deliverables

- Launched experiment
- Notification to broader team of the test launch, including test scope

What to watch out for

- A lack of clarity in how Optimizely tracks data, especially with regard to conflicting results between analytics systems

Keep an eye on health metrics

When you monitor a running experiment, look for issues that may prompt you to stop the test. Once you launch, let stakeholders including your QA team and the team in charge of your production environment that the experiment is live. Share the experiment plan.

Plan to review your Results page a few days after launch to allow a significant number of visitors, at least 5% of your proposed sample size, to see your variations.

Review the following metrics to make sure that your goals are tracking properly and variations are performing in an expected range, based on past results.

First, check if the following goals are within the expected range for your experiment overall.

- revenue
• form submissions
• primary goal
• monitoring goals
• data in third-party integrations
• all other metrics important to your business

If something looks seriously amiss, make sure your experiment is set up correctly. Our standard QA process is a good place to start.

Note:

Often, results fluctuate and a sudden drop in conversions may rise again before your test reaches statistical significance. This variance across time is due to chance -- something like landing on “heads” three times out of 10 in a coin toss. Stats Engine is designed to help you reliably predict future outcomes according to a certain statistical standard. But before you reach statistical significance, how do you know whether a drop in conversions is real - due to an experience that’s bad for your visitors - or just a normal part of running an A/B test?

Check your difference interval. Are you seeing a drop in conversions accompanied by a narrowing difference interval? This combination suggests that Stats Engine is honing in on the likelihood that this variation will convert at the lower conversion rate in the future, with increasing precision. If the scenario continues for days, consider pausing the experiment to assess the impact of continued losses in conversions; and take a moment to evaluate why the variation is losing. The results of this test can help you design better future experiments.

If the conversion rate drops but the difference interval is widening, Stats Engine is projecting a wider range of statistically likely outcomes -- it’s less sure what the conversion rate would be if the test were run again. Remember, if the difference interval straddles zero, Stats Engine has not yet found a statistically significant difference in your conversion rates, positive or negative.

If you have the resources, keep running this experiment to achieve more precise results.

Historical experience can also provide important insight on what to expect from your results.

Once you determine that your overall metrics are healthy, segment your visitors to check key conversion goals for your most important audiences.

SEGMENT YOUR VISITORS

If your experiment is healthy overall, peek into your different visitor segments. Check that conversions haven’t crashed for your most valuable customers.
In Optimizely, you can segment by the following attributes:

- browser type
- source type
- device type (mobile)
- campaign source types (for customers under certain plans)
- custom segments for audiences that are important for your business

Most of the time, a drop in conversions is nothing to worry about -- after all, unexpected results are part of the reason why you’re testing in the first place. The blip may be due to chance, as described in the note above, or it may be an opportunity to learn from the test.

But a steep drop in conversions for just one segment may indicate that you’re delivering a bad or dissonant experience for certain visitors. If that segment of visitors is highly valuable to your business, you may not want to serve a variation that strongly inhibits their conversion.

For example, imagine that returning visitors drive conversions on your site because it takes several visits for customers to feel comfortable purchasing. You segment by new versus returning visitors, and find a steep drop in conversions by returning visitors. At this stage, you might investigate whether there’s a technical issue or if you’re serving a truly bad experience for that valuable segment.

**Pause and re-launch** a running experiment that’s broken or providing an unacceptable experience to important customers.

If your **key metrics and visitor segments are converting as expected**, feel free to let the experiment run according to your test plan.

---

**Pause and re-launch a running experiment**

If you’d like to re-launch a broken experiment or one that’s delivering an unacceptably bad experience, first pause the test in Optimizely.

1. Select the experiment, click the ... icon, and select *Pause*. Once you pause an experiment, visitors won’t see any of the variations.
2. **Duplicate the experiment** to create a new experiment with all the same variations and goals so you can revise the design. If you revise and restart the original experiment, visitors who saw the first experience remain bucketed and will see the same experience when they return (unless they delete their cookies).

3. Use the Editor to remove the variation that you’d like to exclude in the duplicate experiment, and add or modify any other variations.

4. Run the new test. Visitors will be re-bucketed.

That’s it! You are now free to let your test run until the time-to-completion projected by your test plan. When you reach it, you’re ready to interpret and [take action on your results](#).