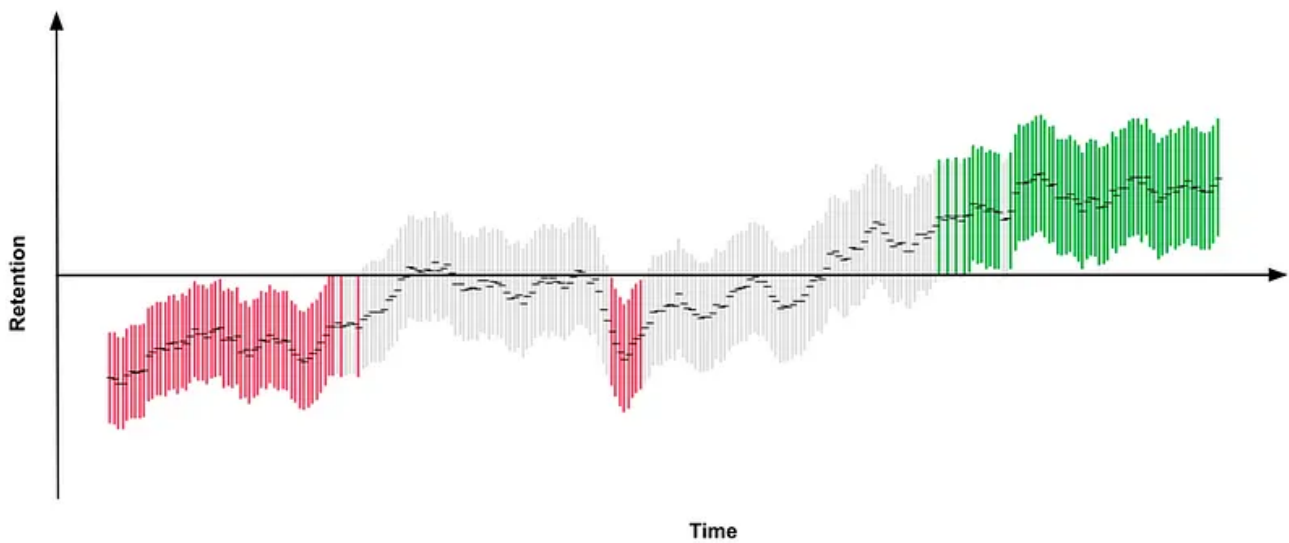


[Open in app](#) ↗[Sign up](#)[Sign in](#)**Medium** Search

Notifications: why less is more — how Facebook has been increasing both user satisfaction and app usage by sending only a few notifications

Analytics at Meta · [Follow](#)

7 min read · Dec 19, 2022



Listen



Share

We are members of the Facebook Notifications Data Science team at Meta. We ran surveys on how users felt about on-site notifications (notifications that show up in the Facebook app when you open the app and create a red badge that counts the number of notifications you have) and found that many users prefer to receive only a few notifications. Based on this finding, we ran an experiment where we sent considerably fewer notifications than usual, and we limited those notifications to ones that we predicted surveyed users would tell us they want the most (e.g. only those with a score of 5, but not 4, on a scale of 1–5), rather than our usual operating model of sending all relevant notifications (either rated 4 or 5 on a scale of 1–5). But in the experiment, we found that people started using the app less frequently. We

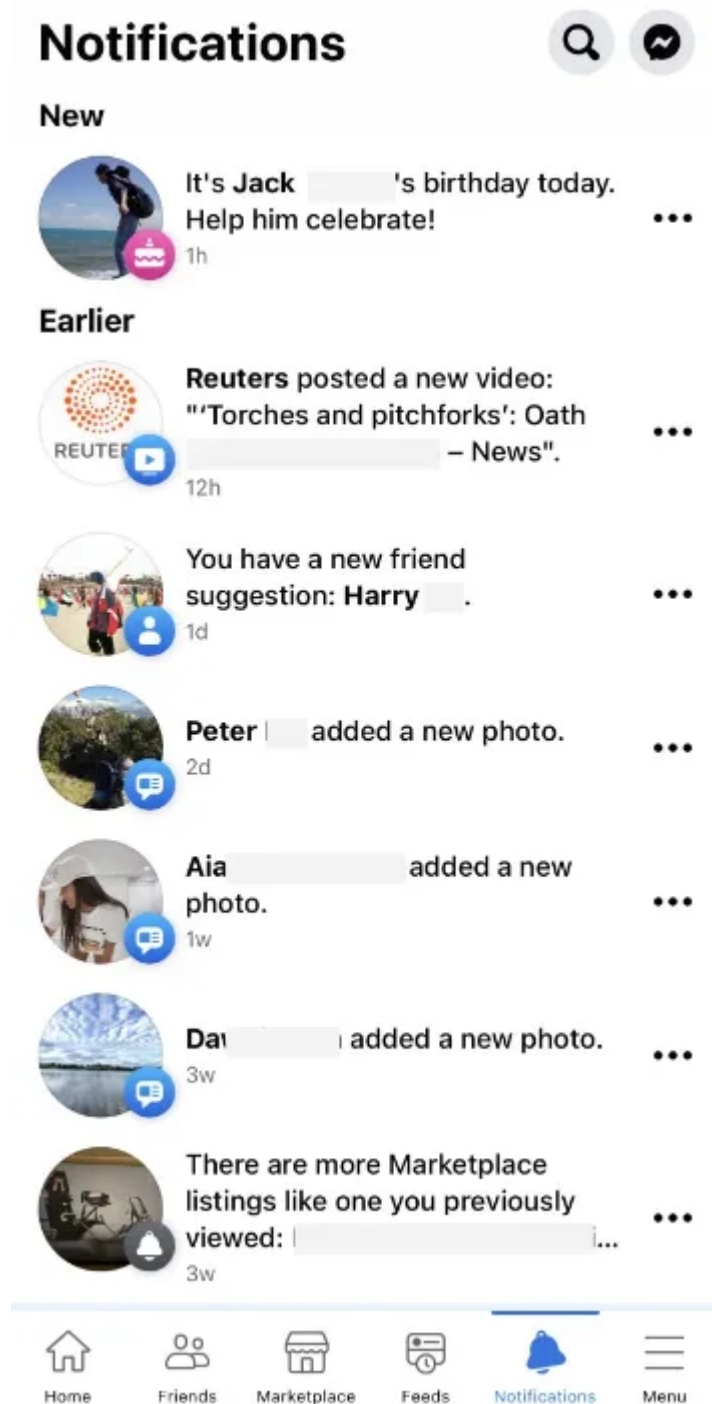
concluded that increased user satisfaction was worth the trade-off of less App usage. Still, it would be better if the improvement of user satisfaction (through surveys) increases the value users are getting out of the product (with App usage as an indicator). After further examination, we found that we can actually increase *both* with a single important data insight: understanding the extent to which long-term effects may differ from short-term effects.

We had a hunch that such a trade-off should be avoidable: if users are more satisfied with their experience, we believe this would be reflected in usage gains too. We decided to keep the experiment of sending only a few notifications running for a year. And lo and behold: little by little Facebook usage started inching back up! After a year we saw that in the fewer notifications experience, users were using Facebook **more — it just took a long time for user behavior to shift and less disruption led to high organic usage, which increased both user satisfaction and app usage.** We wrote this blogpost to socialize this finding as we believe other data science teams outside Facebook could benefit from the same lessons we learnt:

1. Experiments in the *very* long-run can show different results than in the short run.
2. Sending only a few notifications and the most relevant ones (e.g. those predicted to receive a 5/5 rating in the survey mentioned above) can improve user sentiment and usage too in the long run.

What are 'on-site' notifications?

The bell-shaped notifications tab on Facebook's home screen displays on-site notifications. A friend of yours may have posted a picture or someone in a group you are a member of may have left a comment. You can click the notification to view (hopefully!) relevant content



Why does balancing the amount of notifications matter?

To help users get the most value out of our product, one of the major areas we focus on is estimating how many notifications should be sent at the right time. People could miss the content they're interested in if we don't send out enough notifications in time. But on the other hand, if we send too many notifications (or notify users with less relevant updates), even though users may visit Facebook more frequently in the near term, they may also perceive notifications as spam, develop a negative opinion on them or opt out from receiving notifications with the control options they have, which might make them miss relevant content in the future. We don't want to optimize our product towards this unsustainable short-term manner; rather,

we think that long-term value that guarantees people have a worthwhile experience is what we should be working for.

How do we measure this?

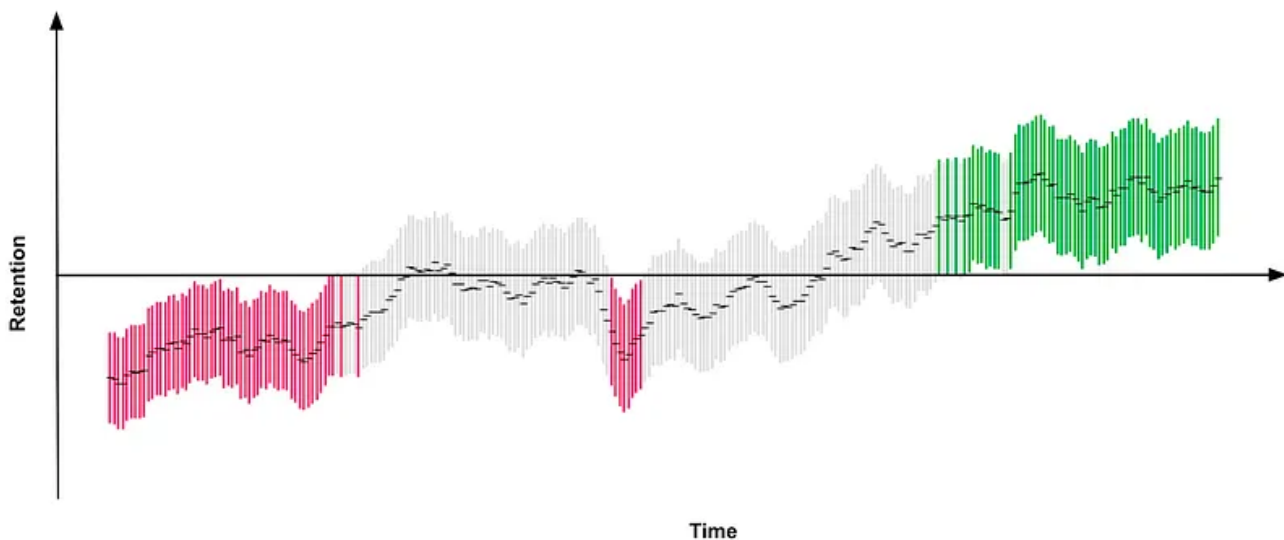
Through user research, we learned that our users want to receive only the most relevant notifications (based on user surveys where users rated the relevance of actual notifications they received). So we set up an A/B test to compare what happens when users send only notifications we predict to be very relevant vs. a wider array of relevant notifications (e.g. we predict users would rate them 5 on a 1-to-5 scale instead of 4, in neither test group did we send notifications we predict to be irrelevant).

A/B testing is a widely used technique in the industry, to determine the causal relationships between the number of notifications and the targeted outcome (e.g. product usage, user satisfaction level). Therefore, we conducted a series of tests where we sent different amounts of relevant notifications to randomly allocated groups of users. We then measured various outcomes ranging from product usage to user satisfaction indicating metrics.

What did our measurement teach us?

Even if users claim they don't enjoy getting too many notifications, the usual A/B testing results do commonly suggest that sending only a few notifications would immediately result in reduced product visits. One drawback of A/B testing is that the experiment duration is typically brief (a few weeks or months, for example), making it impossible to detect if the result is purely due to novelty effect, or real change of user satisfaction with our product. We chose to extend the fewer notifications experiment for many months and even years, in order to learn about the long-term effects. Of course, it will take a much longer time for us to reach a conclusion, but that's what we should do to learn important long-term learnings, which can be very different from the short term learnings.

We discovered that the initial loss of visitation from fewer notifications gradually recovered over time, and after an extended period, it had fully recovered and even turned out to be a gain. In other words, long-term effects may be different from short-term effects, or even the opposite. It may take time for people to adapt to the change. Our results suggest people will eventually return and become even more engaged with Facebook if the change truly improves their experience.



OK, so how do I set up a long-term experiment?

Long-term experiments need to be set up differently to short term ones.

First, if you run an experiment for a year, the difference between the initial and year-long effects might not be causally due to time differences. For example, new people might have signed up to Facebook over the last year, or the content might be different now than in the start of the experiment. To tell apart these “ecosystem-level” changes from true short-run vs. long-run differences, we conducted a new replicated experiment a few months after the first one started, and let it run for a while, and then compared the initial trend to the original experiment to see if it coincided with the replicated one. The likelihood of the launch reproducing the long-term effects of the initial experiment increases if the initial trend is in agreement. After confirming that the initial effects were in line, we launched our product change to only send the few very relevant notifications to our users.

But what if I can't wait?

A scientific method for determining the difference between short-term and long-term effects is to conduct long-term experiments. You will have to wait a very long time for the results, which is obviously a disadvantage of this method. However, some experiments are difficult or undesirable to run for a long time, and you may not always have the luxury of time for waiting. There are still a few alternative things you could do to study long-run effects.

First, you could assess short-run experiments using the knowledge gained from earlier long-run experiments. For instance, we discovered from the experiment with only a few notifications that there would be a short-term loss in visitation but a long-

term gain. We naturally wonder if experiments with immediate visitation gains that involve sending more notifications will eventually lose their immediate gains or even result in long-term visitation losses. In order to determine whether the gain has a downward trend, we will be cautious before declaring the gain and require such experiments to run for a little longer.

Additionally, you could create a long-term proxy based on experiments (<https://medium.com/meta-analytics/estimating-the-long-run-value-we-give-to-our-users-through-experiment-meta-analysis-6ddb9073b29b>) and use it to calculate the long-term impact of short-run experiments. The basic idea is to first conduct a number of long-run experiments, after which the experiment data will be used to fit a regression to determine the coefficients of various treatments toward long-term effects. You could then predict the long-term effects for a short-run experiment using the coefficients from the regression. For instance for outcome Y_i (eg user satisfaction measured on a year-long horizon) you can predict it with short-run observable metrics x_1 , x_2 and x_k such as the number of notifications sent, and fit an artificial intelligence model to derive coefficients leading to a reusable formula such as:

$$Y_i \sim coeff_1 * x_{1i} + coeff_2 * x_{2i} * \dots * coeff_k * x_{ki},$$

You can then use the formula derived in short run experiments. Furthermore, it's critical to update the formula on a regular basis, such as once every six months.

Only fools rush in...

In conclusion, one typical A/B testing pitfall is that people may make decisions based on the results of short-run experiments, while some product changes may have quite different long-term effects. Beware of novelty effects, and measure them as best you can. If it is difficult to run experiments for a long time, you may try extrapolate learnings from past long-run tests or create an experiment-based proxy to assess short-run experiments. Your wait is worth it and you will make better decisions with longer experiments and AI models based on them.

Authors: Weijun C., Yan Q., Yuwen Z., Christina B., Akos L., Harivardan J.



Follow



Written by Analytics at Meta

9.5K Followers

The mission that unites Meta Analytics is to “drive better outcomes using data as a voice for our communities.”
