**PRACTITIONER'S GUIDE**                    **2021** MARCH

# Designing Adaptive Experiments

STANFORD GRADUATE SCHOOL OF
BUSINESS | Golub Capital
Social Impact Lab

Northwestern | Kellogg
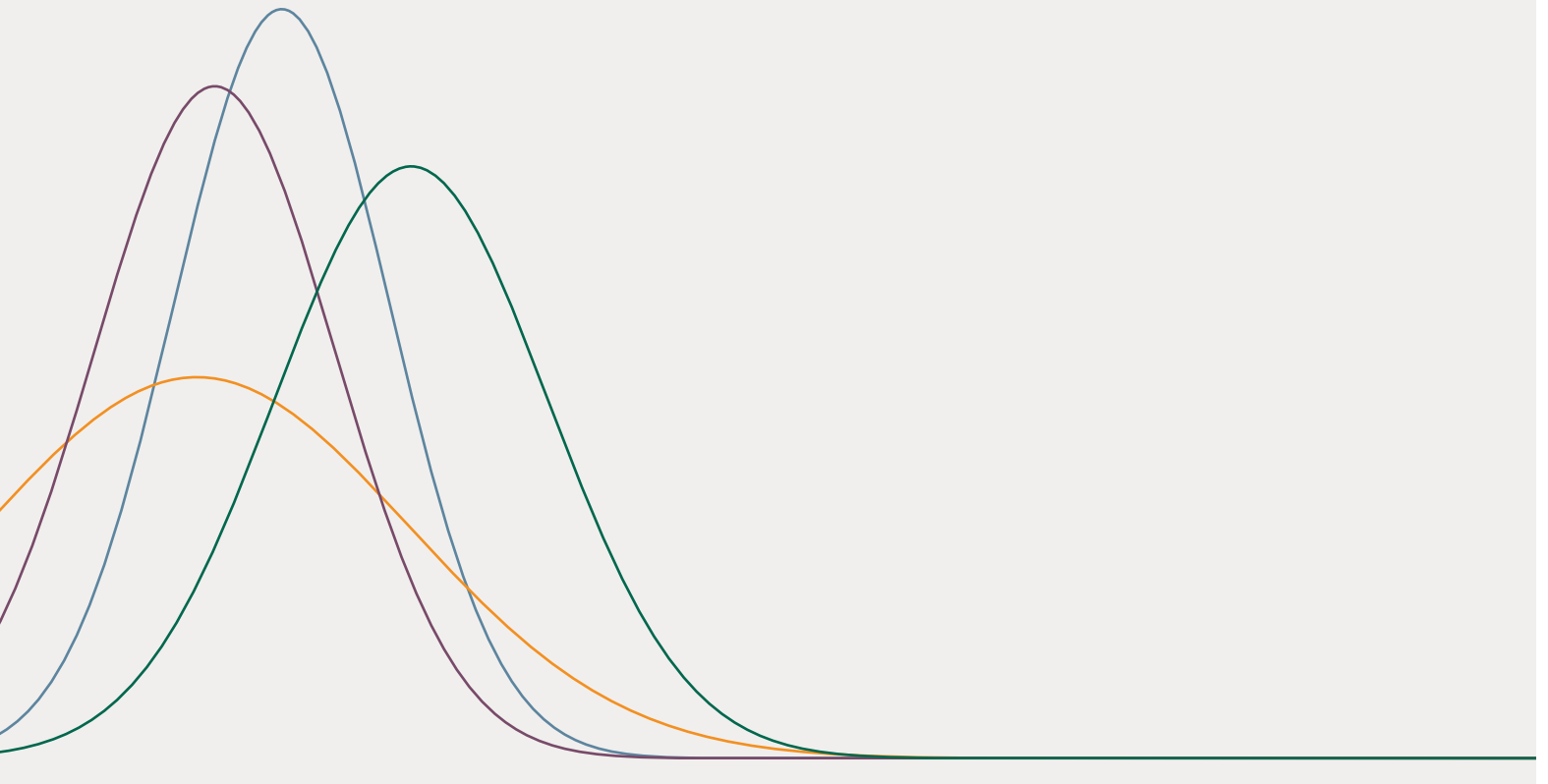
Adaptive experiments present a unique opportunity to more rapidly learn which of many treatments work best, evaluate multiple hypotheses, and optimize for several objectives. For example, they can be used to pilot a large number of potential treatments when the researcher does not have strong hypotheses about what works and why; the data can then be used to narrow down a set of alternatives for further development, hypothesis testing, and evaluation. However, adaptive experiments may not be appropriate in all cases, as they create complexity and may require additional resources for implementation. This guide explains what adaptive experiments are, when they can be beneficial, and their limitations. It also offers insights into the questions to ask when considering running adaptive experiments on technology platforms.

# Introduction

## WHAT ARE *ADAPTIVE* EXPERIMENTS?

This section introduces some basics of adaptive experimental design.

To start, consider some definitions. The purpose of an experiment is to learn about the effect of particular interventions, called *treatments* or *arms*, on some *outcome* of interest. This definition is very broad and encompasses many situations. For example, in the context of a clinical trial, treatments may be different drugs, and the outcome is a measure of health. For a video streaming service, treatments may be different recommendation algorithms, and the outcome is whether customers watch the recommended video. In an online learning platform, treatments can be different teaching tools, and the outcome is student learning or test scores.

Let's focus on one specific example, where the goal is to select a website design. The treatments are different color schemes, and the outcome is some measure of visitors' engagement with the website, such as their time spent on it. We represent these color schemes as the four colored squares in **FIGURE 1**. The circles represent our *experimental budget*, which in our setting corresponds to the next 100 visitors to the website. In the experiment, we will assign each visitor to a particular color scheme. At the end, we will tally up how much each visitor engaged with the website to determine which color worked best.
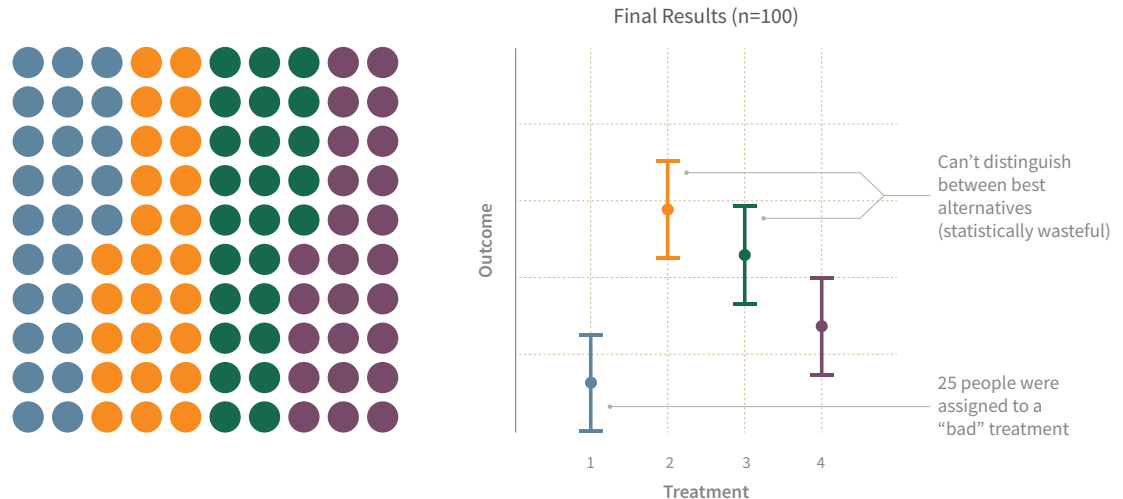
Figure 1
Experiment Setup

Treatments or Arms

Experimental Budget (n=100)
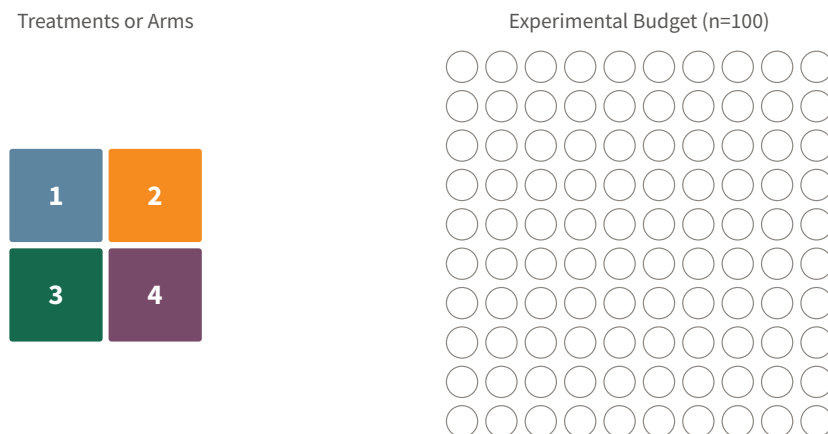
The crucial property that nonadaptive experiments have is that the fraction of observations assigned to each treatment is set before the experiment starts. Continuing our online experiment example, in a nonadaptive experiment, we would need to select the number of users who will observe each scheme before the experiment starts.

## Illustration of Nonadaptive Experiment Results



The left panel of **FIGURE 2** demonstrates treatment assignment in our proposed experiment. Note that each arm received an equal share of observations (25 people were assigned to each treatment). On the right panel of **FIGURE 2**, we present illustrative results. The 95% confidence intervals represent uncertainty around the estimate of the average performance of each arm with respect to our outcome of interest.

At this point, we should step back and consider whether this experiment was successful at helping us determine the optimal website design. In fact, we find that we cannot statistically distinguish between the orange and green arms, so we fail to identify a single best treatment. There may be two sources of inefficiency here. First, if we were interested in learning more about the best-performing treatments, then in hindsight we would have preferred to collect more observations from the orange and green treatment arms — that is, to have routed more users to those colors. The second source of inefficiency is that many observations were assigned to the purple and blue arms, which, as we learned only after the experiment was over, are low-performing arms. The

fact that many observations received poor-performing arms is also a source of inefficiency. In our example, we missed the opportunity to engage those website visitors, which could be costly. In other, more serious, cases, such as drug trials, it could also be considered unethical.

If we had the benefit of hindsight, we could have designed the experiment with fewer inefficiencies — for example, by assigning fewer observations to poor-performing arms. Of course, in reality, we don't have that information at the beginning of the experiment. The goal of adaptive experiments is to minimize these inefficiencies. The idea is to set a particular objective and then collect data to maximize that objective during the experiment. We do this by sequentially changing the proportion of observations assigned to each arm as the data comes in and we gather more information about the properties of each arm.

**FIGURE 3** shows an example of an adaptive experiment. Let's say that our objective is to route online traffic to the version of our website that leads to the most engagement during the experiment. Of course, at the

beginning of the experiment, we have no data on any treatment, so we cannot know which ones lead to higher or lower engagement. Therefore, we begin by assigning them with equal probability. However, after collecting a small wave of data (say n=20 in the example), we estimate intermediate results and get some idea of the performance of each arm. These intermediate results inform how the next wave of data should be collected.

That is, we assign the arms that seem more promising more often, according to the objective we set out. (We'll discuss the exact mechanism through which assignments are decided in the next section.) Adaptive experiments proceed sequentially in this manner, updating how frequently we assign each treatment based on past data.
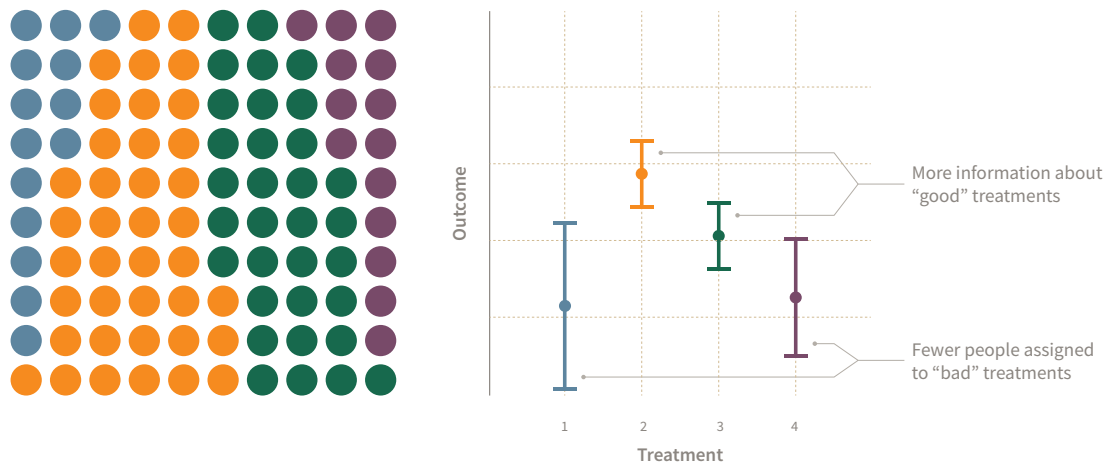
## Illustration of an Adaptive Experiment
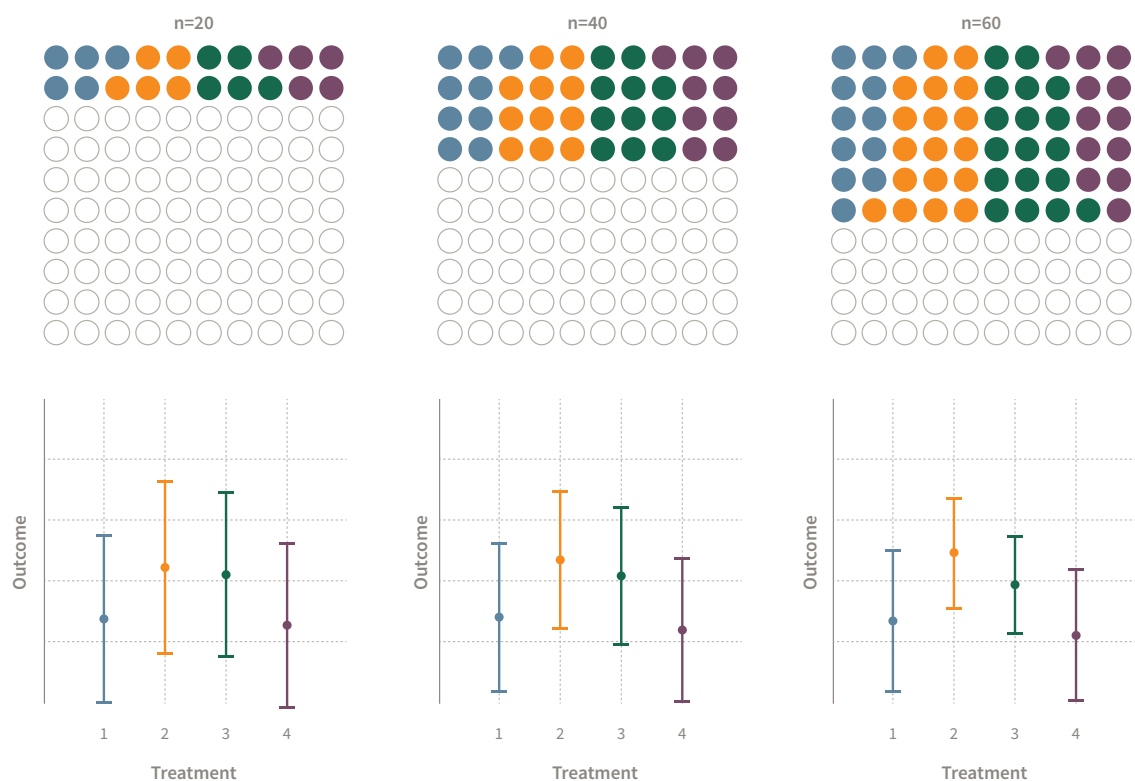


FIGURE 4 shows an example of the results at the end of this particular adaptive experiment. Compared to the nonadaptive experiment in FIGURE 2, our adaptive experiment should assign high-performing arms more often and low-performing arms less often. This allocation of observations away from low-performing to high-performing arms has the benefit of reducing the type of inefficiencies mentioned previously.

Figure 4

## Illustrative Adaptive Experiment Result



Final Results (n=100)

More information about "good" treatments

Fewer people assigned to "bad" treatments

In considering how to assign treatments, experimental designers must balance two opposing forces: the need to learn about the value of each treatment arm, and the need to assign arms optimally according to a particular objective. This problem is often called the *exploration-exploitation trade-off*.

We'll close this section by emphasizing that, although here we posited a particular objective

(minimizing the number of observations assigned to low-performing arms during the experiment), adaptive experiments can be used for several different objectives. In the next section, we delve a little deeper into the adaptive experiment design literature and discuss concrete examples of experiment design that optimize the exploration-exploitation trade-off for different objectives.

# Adaptive Experiment Objectives

In the previous section, we mentioned that the goal of adaptive experimental design is to collect data with a particular objective in mind. In this section, we discuss some examples of different objectives and illustrate specific algorithms. We emphasize that the goal here is only to give the reader a high-level, nonmathematical overview of the technique.

Perhaps the most studied problem in adaptive experimental design is maximizing outcomes during the course of the experiment. This is often called **cumulative regret minimization**. *Regret* is a technical — but widely

used — term that is defined as the difference between the average value of the assigned treatment and the average value of the best treatment. For example, if Treatment 1 obtains an average outcome of 3 units but the best treatment would be able to obtain an average outcome of 8 units, then by assigning Treatment 1 we incur a regret of 8-3=5 units. Of course, minimizing regret implies maximizing rewards, but it's often more convenient to reason in terms of regret. Cumulative regret minimization algorithms represent much of what is known as **multi-armed bandit algorithms**.

However, multi-armed bandit algorithms are not appropriate for every experiment. Sometimes, we simply want to find a good treatment at the end of the experiment, regardless of what happens during it. Depending on how exactly the problem is set up, this problem is called **best-arm identification** or **welfare maximization**. Finally, we may instead want to collect data to test a particular data-driven hypothesis.

Although the theoretical literature on adaptive experimental design usually focuses on one of these objectives at a time, in practice, experiments often have multiple objectives. For instance, industry partners might be particularly concerned about their reputation as an engaging platform. A bad user experience could lead to disengagement and bad press, and therefore the company might want to design an experiment to achieve sufficiently high outcomes during the experiment while also finding a good (or best) treatment after the experiment. More generally, there may be different objectives based on short-term and long-term goals for the research and the product. Researchers' incentives can also differ from those of the industry partner. Whereas researchers may be interested in estimating the effectiveness of particular "losing" treatments, even if null or negative, the partner may care only about learning what treatment works best. These varied objectives of adaptive experiments mean it is imperative to discuss and align on the particular objectives and design of the experiment.

In the following sections, we consider several examples of different objectives, and we analyze how experimental design can be tailored to those objectives.

## MAXIMIZING OUTCOMES DURING THE EXPERIMENT

A large portion of the literature on adaptive experimentation deals with the objective of maximizing expected outcomes during the experiment, meaning assigning fewer people to poorly performing arms.

This might be the primary objective in instances when researchers are concerned that poor arms might actually contribute to harmful outcomes. For example, in an online study of interventions to curb sharing of misinformation, it is conceivable that a treatment could increase sharing of misinformation. We would want to know if a treatment contributes to this kind of bad outcome and minimize this during the experiment, rather than waiting until the end. In such a case, the researcher might balance avoiding harm during the experiment with gaining insight that can be used in the future; below, we discuss balancing objectives.

To see a scenario where maximizing outcomes during the experiment is the main objective, consider a scenario where a firm is trying to maximize revenue, profit, or social impact. For example, a charity might want to encourage individuals to donate each time they arrive at the charity's website. The charity wishes to learn which call to action maximizes the expected donation per visit. Ineffective calls to donate represent a missed opportunity for the charity, and ongoing experimentation has a cost relative to sticking with the call to action that looks best based on historical data. If experimentation is going to be a long-run, ongoing process (that is, the experiment never ends, perhaps because new ideas are introduced over time), then the charity should try to maximize donations during the experiment.

Two common classes of algorithms used in adaptive experiments to maximize outcomes during the experiment are **Upper Confidence Bound** methods and **Thompson sampling** methods.

**Upper Confidence Bound (UCB) algorithms**: After each observation, we compute a particular kind of confidence interval around the sample mean of the outcomes collected so far. Then we select the arm whose confidence interval upper endpoint is highest. **FIGURE 5** shows an example. Note that even though Treatment 2 (orange) seems to have the highest estimated sample mean, a UCB algorithm would select

Treatment 3 (green) in this setting since the upper endpoint of its confidence interval is the highest. This forces us to collect more data about treatments that we are uncertain about, but in a manner that still favors the most promising ones – a heuristic often called "optimism in the face of uncertainty."

## Selecting an Arm Via the "Upper Confidence Bound" (UCB) Heuristic
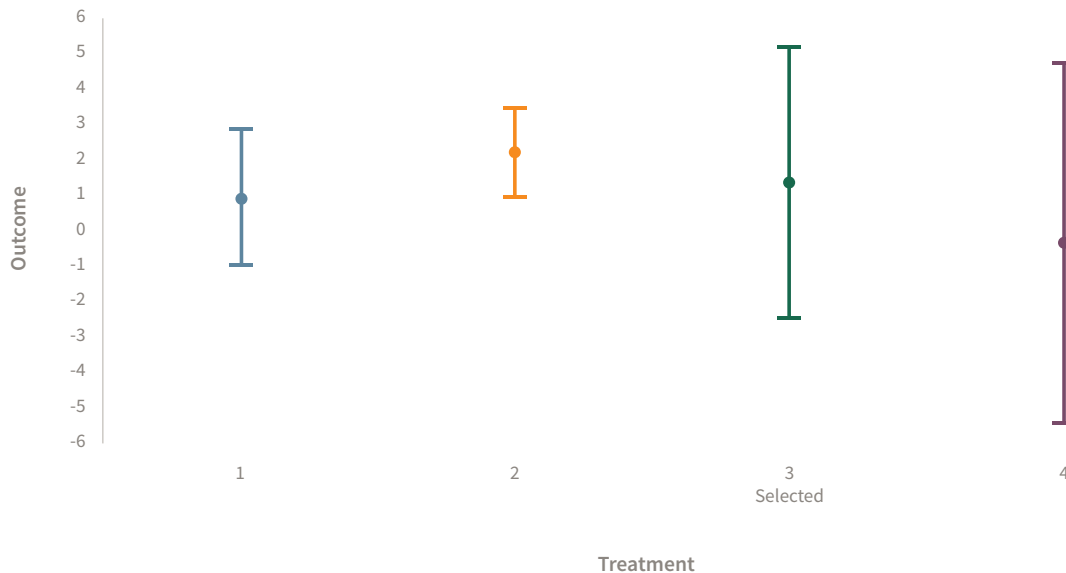


**Thompson sampling algorithms**: In this class of algorithms, we maintain a Bayesian probabilistic model of which treatment is best and assign treatments according to this model. For example, if at a certain point in the experiment the model suggests that there's a 40% chance that Treatment 1 is better than any other treatment on average, then Treatment 1 is assigned with 40% probability. As more data is collected, the probabilistic model gets updated, and so do the treatment assignment probabilities.

This evolution is illustrated in **FIGURE 6**. At the beginning of the experiment, there isn't any data, so we begin by assigning all treatments with equal probability (25%). However, after the first wave of 20 observations, Treatments 2 and 3 may seem more promising, which leads us to assign them with higher probability (about 50% and 30%, respectively). After that, as more data are collected, the algorithm is better able to distinguish between treatment values and ends up heavily favoring Treatment 2 by the end.

**Figure 6**

Treatment Selection Probabilities Computed via the
Thompson Sampling Heuristic as the Experiment Progresses
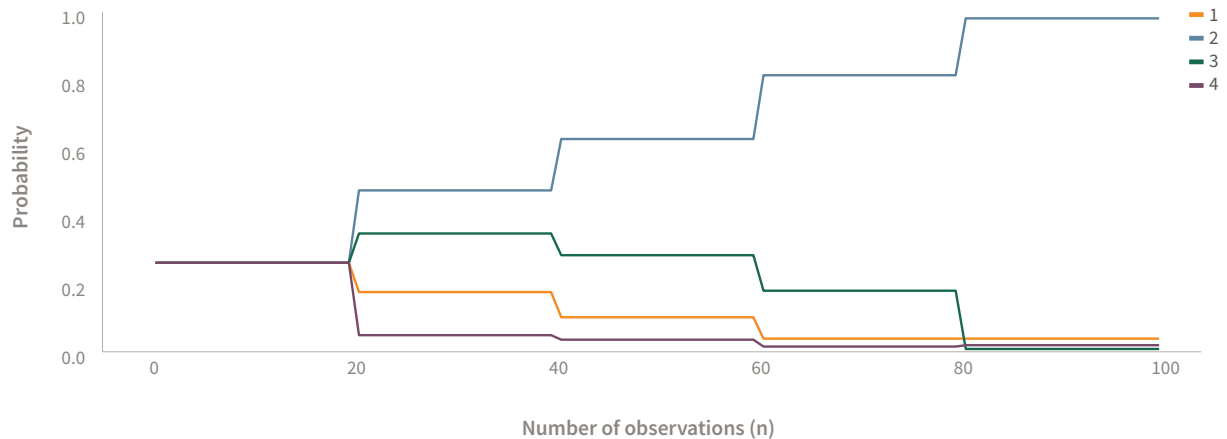


## FINDING A GOOD TREATMENT AT THE END OF THE EXPERIMENT

The previous algorithms have been shown to be effective at maximizing outcomes during the experiment. What if our objective was instead to find a good treatment at the end of the experiment — regardless of what happens during it? For example, an online merchant might want to test several different scripts that ask their users to donate $1 to charity during their online checkout. In this case, the objective is to identify the best script at the end of the experiment that leads to the greatest number of donations. Here, assigning users to poorly performing arms does not have as negative welfare implications as it might in other contexts — for instance, where seeing a bad arm might increase the spread of misinformation. So in this case, we are comfortable with assigning some people to poor-performing treatments during the experiment and learning the best overall at the end. This problem is commonly named *pure exploration*.

It turns out that the algorithms mentioned in the previous section may not be optimal for this alternative objective, but there exist a few heuristics that try to tackle this problem. For example, **Exploration sampling** (Kasy and Sautmann, 2021) proposes a modification of Thompson sampling probabilities to make them less aggressive, forcing additional exploration but also increasing the expected value of the arm selected at the end of the experiment. For a review of other algorithms in this literature, see Jamieson (2015).

## OTHER OBJECTIVES

The objectives of an adaptive experimental design can also be more of a statistical nature, such as maximizing the number of significant findings (van der Laan, 2008). Also, experiments often have a combination of objectives. For example, we may want to not only find the best treatment at the end of the experiment but also test whether the treatment we selected is better than some status-quo control treatment. In that case, it is important to have collected enough observations for the control treatment, even if its performance is poor. For such situations, the algorithms above need to be modified further, for example by imposing a lower bound on assignment probabilities that ensures that relevant treatments we care about estimating are always selected with a sufficiently high probability.

## EXTENSIONS: PERSONALIZATION

There also exist extensions of the algorithms above that allow for personalization, by changing the probabilities that each individual will be assigned to different treatments depending on the individual's observable characteristics. In the literature, observable covariates are often called *contexts*, and algorithms that use them are called **contextual adaptive experimental designs**. Contextualization can be simple or complex: We may simply allow for different assignment probabilities for different pre-specified subgroups; we may allow the subgroups to be data-driven and evolve over time, or we may use a full-blown statistical model of outcomes as a function of contexts and treatments.

Contextual adaptive experiments are useful when we expect that particular treatments will be best suited for different types of people or subgroups. For example, in a study designed to test interventions to curb the spread of online misinformation, we might expect that a video training that provides tips to spot fake news would be the best intervention to assign to social media users with low digital literacy. By contrast, an intervention designed to nudge users to consider the veracity of a story might be best at reducing misinformation sharing among social media users who have high digital literacy skills and make quick intuitive decisions, rather than deliberating (Offer-Westort, Rosenzweig and Athey, 2021).

On the other hand, if we anticipate that a treatment arm that is best for certain people is harmful to others, accounting for context in an adaptive experiment may be critical for ethical reasons, not just for helping to optimize experiment design. Continuing our online misinformation example, if we anticipate that a particular treatment will work well for older users but might increase sharing of misinformation among younger users, we would want to plan for this in the design of the experiment. Assigning treatment by context is one way to deal with this kind of situation.

Allowing for personalization is not always advantageous. To see why, imagine an adaptive experiment whose contextualization is based on membership in one of four subgroups. Now consider the two scenarios presented in **FIGURES 7** and **8**. In **FIGURE 7**, the data analysis shows that each subgroup seems to prefer a different arm. This information would be obscured if we analyzed the aggregate data without considering the subgroups, so personalization is indeed relevant here. However, **FIGURE 8** presents a different situation, where all groups prefer the same arm. If this is the true data-generating process, then personalization may in fact delay learning the optimal treatment, since we are needlessly breaking down our data into subsets, which also reduces our statistical power.

Figure 7

## Illustration of Strong Heterogeneity

A scenario in which it's possible to detect that each subgroup prefers a different arm. In this case, personalization should help.
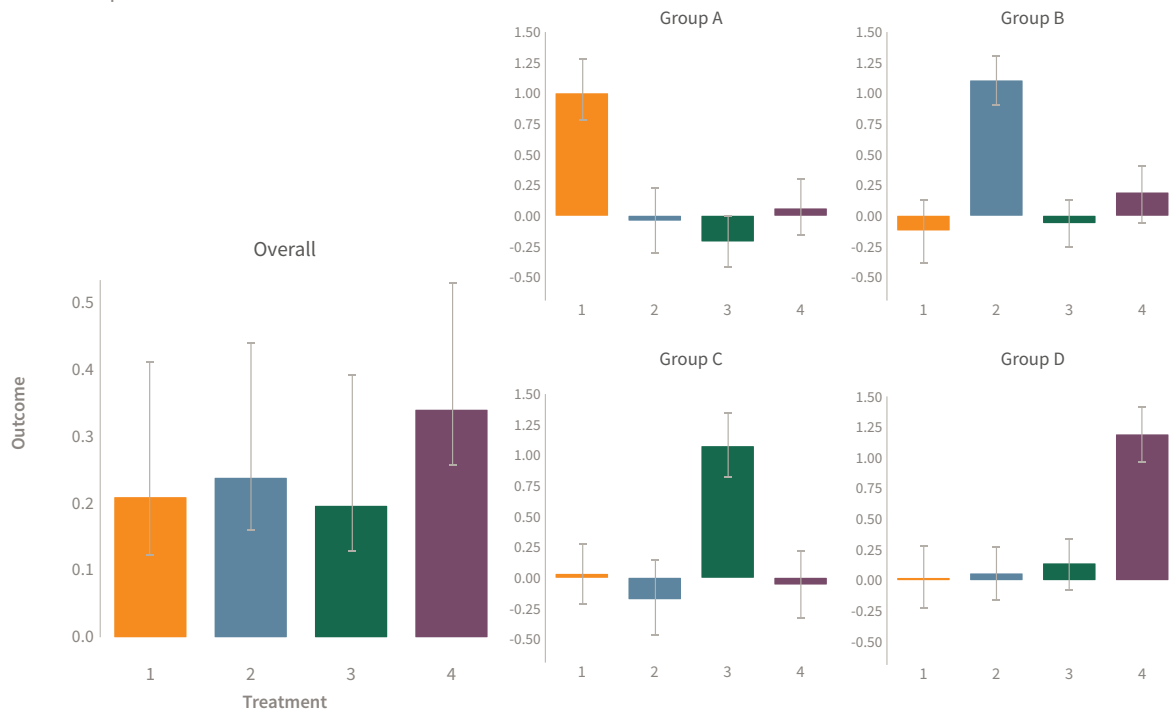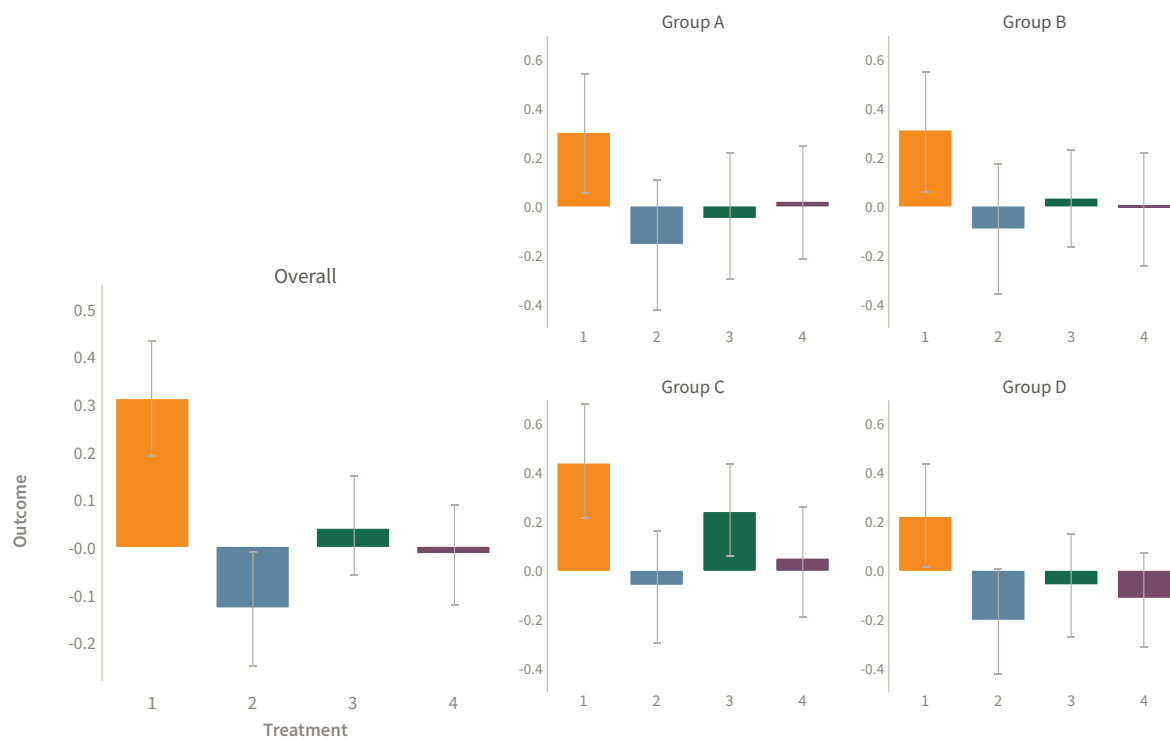


Figure 8

## Illustration of No Heterogeneity

Every group prefers the same arm. In this case, personalization during data collection will only delay detecting the right arm.

# Benefits and Drawbacks of Adaptive Experiments

## BENEFITS

When designed well and specified appropriately, adaptive experiments can offer many advantages relative to nonadaptive ones.

First, adaptive experiments can provide efficiency gains: They help us get the right treatment to the right people more quickly. This happens more quickly since we're not waiting until the end of the experiment to know the "answer." Rather, we send more traffic to better-performing treatments throughout the experiment. This not only improves efficiency from a research perspective, but in certain situations it can also be more ethical, cost-effective, and welfare-enhancing.

Second, adaptive experiments can help save human time. If researchers must narrow down to two or three treatments in advance, as opposed to testing all proposed treatments, they may spend substantial time considering which is likely to be best or running pilot studies. In contrast, there is a smaller cost to incorporating additional treatments with adaptive experiments, since the experiment will adapt and allocate more units to the best treatments during the experiment. Relatedly, it is also possible to use adaptive experiments for prototyping. For example, running a pilot experiment with many arms in an adaptive manner can help to remove uninteresting arms quickly and identify the smaller subset of better-performing arms to test in a traditional nonadaptive experiment.

Third, adaptive experiments can optimize for multiple objectives. Specifically, the objective of the experiment can incorporate costs that are associated with the experiment (e.g., transaction costs or frictions, such as users losing attention or leaving the platform before completing the call to action). For example, the online merchant interested in finding the best appeal to donate $1 to charity at checkout might optimize for the highest number of donations while also minimizing the number of abandoned purchases.

## DRAWBACKS

On the other hand, designing, running, and analyzing an experiment can be more difficult for an adaptive experiment than for a nonadaptive one. Many of the usual concerns associated with a traditional experiment are exacerbated in adaptive designs. Here we touch on a few key considerations.

First, in general, planning becomes more complex. The experimenters need to decide on a particular objective, which may require balancing the interests of multiple stakeholders. Next, they must decide on an appropriate algorithm that satisfies this objective. Some algorithms have tuning parameters whose configuration may require additional work, such as simulations, debugging, etc. Moreover, as we have seen in the previous section, deciding on the level of personalization of the adaptive design requires careful consideration. Resources for planning an adaptive experiment are also harder to find and often require expertise.

Second, there are engineering costs in setting up adaptive experiments. The magnitude of these costs depends on the complexity of the experiment, so it's advisable to keep it simple, at first. For example, an experimentation team that already has an A/B test pipeline can begin by doing sequential A/B tests, with earlier experiments informing later ones. As the team gathers expertise, they can automate the updating of treatment assignment probabilities. Finally, on a longer timeline, personalization can be introduced if its statistical and welfare benefits outweigh the costs of collecting additional contextual data and computing personalized assignment probabilities.

Finally, once the data are collected, the statistical analysis of adaptive experiments requires more sophisticated tools than nonadaptive ones. Statistical techniques that do not take adaptivity into consideration — including usual t-tests — can be biased or have incorrect confidence intervals (see

e.g., Hadad, Hirshberg, Zhan, Wager and Athey, 2021). It's important that the experimentation team be able to understand and apply the appropriate statistical methods. In addition, communicating the analysis of the results to partners may require more nuance.

# Conclusion

We hope this guide provides a high-level overview of adaptive experiments for researchers and businesses interested in exploring these techniques. To summarize, for anyone thinking about conducting an adaptive experiment, we propose the following three steps:

1. Think about whether this project would benefit from adaptivity by identifying and clearly articulating your specific objective(s).

2. Consider whether you anticipate strong benefits to personalization and therefore a contextual adaptive experiment might make sense. Will different arms be best for different subgroups?

3. Discuss the drawbacks of adaptive designs, including more complicated planning, pipelines, and analysis, as well as challenging communication.

Readers looking for additional details on how to run adaptive experiments and how to make some of the decisions we outlined above should check out the resources below and keep an eye on our website for newly developed online tools and tutorials.

# References

*Adaptive & Iterative Experimentation*. (n.d.). Stanford Graduate School of Business. Retrieved March 29, 2021, from *https://www.gsb.stanford.edu/faculty-research/centers-initiatives/sil/research/methods/adaptive-iterative-experimentation*

Dimakopoulou, M., Zhou, Z., Athey, S., & Imbens, G. (2018). Estimation Considerations in Contextual Bandits. *ArXiv:1711.07077 [Cs, Econ, Stat]*. *http://arxiv.org/abs/1711.07077*

Dimakopoulou, M., Zhou, Z., Athey, S., & Imbens, G. (2018a). Balanced Linear Contextual Bandits. *ArXiv:1812.06227 [Cs, Stat]*. *http://arxiv.org/abs/1812.06227*

Hadad, V., Hirshberg, D. A., Zhan, R., Wager, S., & Athey, S. (2021). Confidence intervals for policy evaluation in adaptive experiments. *PNAS*, April 13, 2021 118 (15) e2014602118 *https://www.pnas.org/content/118/15/e2014602118*

Jamieson, K. (2015). The Analysis of Adaptive Data Collection Methods for Machine Learning. Ph.D. thesis, University of Wisconsin-Madison. *https://homes.cs.washington.edu/~jamieson/resources/kevinJamieson_Dissertation.pdf*

Kasy, M., & Sautmann. A., (2021). Adaptive treatment assignment in experiments for policy choice. *Econometrica*, 89(1), 113-132. *https://www.econometricsociety.org/publications/econometrica/2021/01/01/adaptive-treatment-assignment-experiments-policy-choice*

Offer-Westort, M., Rosenzweig, L., & Athey, S., (2021). Optimal Policies to Battle the Coronavirus "Infodemic" Among Social Media Users in Sub-Saharan Africa. *https://osf.io/3ukw9/*

van der Laan, M. J. (2008). The Construction and Analysis of Adaptive Group Sequential Designs. U.C. Berkeley Division of Biostatistics Working Paper Series 232. *https://biostats.bepress.com/ucbbiostat/paper232/*

Xiong, R., Athey, S., Bayati, M., & Imbens, G. (2020). Optimal Experimental Design for Staggered Rollouts. *ArXiv:1911.03764 [Econ, Stat]*. *http://arxiv.org/abs/1911.03764*

# About Us

## STANFORD GOLUB CAPITAL SOCIAL IMPACT LAB

The Golub Capital Social Impact Lab is a lab and research initiative of affiliated academics, staff, researchers, and students who are passionate about applying technology and social science research to improve the effectiveness of leading social sector organizations. The lab is based out of the Stanford Graduate School of Business and collaborates with a wide range of organizations, from large firms to smaller startups, for-profits to nonprofits, and NGOs to governments, to foster collaborative original research.

LEARN MORE: *www.gsb.stanford.edu/ faculty-research/centers-initiatives/sil*

## GLOBAL POVERTY RESEARCH LAB

The Global Poverty Research Lab is an academic hub for empirical development economics at Northwestern University. The lab aims to use empirical evidence and interdisciplinary engagement to understand drivers, consequences, and solutions to poverty issues around the world. It works to address the challenges of overcoming poverty and improving well-being in the developing world. The lab's goal is accomplished by supporting graduate student and faculty research, providing undergraduate research opportunities, coordinating a seminar series, and promoting policy engagement.

## BILL & MELINDA GATES FOUNDATION

The Bill & Melinda Gates Foundation is a private foundation founded by Bill and Melinda Gates. Its mission is to create a world where every person has the opportunity to live a healthy, productive life.

## STANFORD GRADUATE SCHOOL OF BUSINESS

Stanford Graduate School of Business has established itself as a global leader in management education through educational programs designed to develop insightful, principled, global leaders. Stanford GSB supports faculty research, curriculum development, and interaction among academic disciplines.

LEARN MORE: *www.gsb.stanford.edu*

## Contributors

### CO-AUTHORS

**VITOR HADAD** Postdoctoral Scholar
Stanford Golub Capital Social Impact Lab

**LEAH R. ROSENZWEIG** Postdoctoral Scholar
Stanford Golub Capital Social Impact Lab

### FACULTY

**SUSAN ATHEY**
The Economics of Technology Professor
Faculty Director, Golub Capital Social Impact Lab,
   Stanford Graduate School of Business

**DEAN KARLAN**
Professor of Economics and Finance
Co-Director, Global Poverty Research Lab,
   Kellogg School of Management,
   Northwestern University