



CLIMATE
ADVOCACY
LAB

Testing & Evaluation



Last updated 6/22/20

Why Run Tests?

**“Data is how you
listen to people
at scale”**

—FB analytics lead

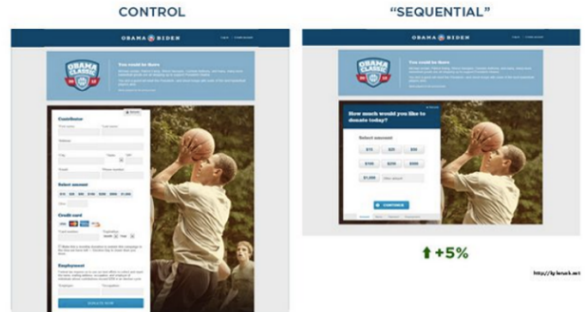
Why do we test?

- We want to know if **x** (our program) **causes y** (action on climate and energy issues!)
- We are not always great at guessing 😊
- Our work is **too important** and **too urgent** to spend time and resources in pursuit of strategies that are not moving us toward winning more policy victories and building more political power!



A/B testing: Obama 2012

Constantly looking for improvements, large or small, in every aspect of our digital operation



The testing mantra: test every element

Running tests in the background via personalized content

- Example: 300+ buckets for the "ask amount"

From: Barack Obama <info@barackobama.com>
Subject: My last campaign

David --

In a few days, I'll be hitting the trail for my last campaign.

Everything we've accomplished in the past three years -- and our chance to do so much more -- is on the line.

What we do today will be a measure of whether or not we're ready to fight for it.

[Donate \\$61 or whatever you can before tonight's fundraising deadline.](#)

By pitching in before midnight, you'll automatically be in the running to join me and George Clooney at his place on May 10th.

It's not often I can get away from work, so I look forward to spending a fun evening in L.A. with a couple supporters like you.

“We basically found our guts were worthless.”

Why does identifying best-performing content and evaluating your programs matter?

Because what you learn can have big consequences



One progressive testing firm found top performing ads often twice to 2x more effective than media ads



Another's testing has found that about 1 in 5 ads backfire (persuading audiences in the *opposite* direction)



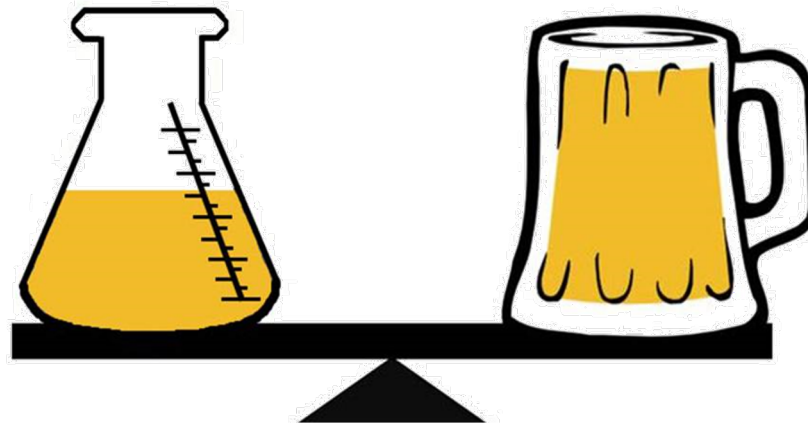
A review of tests in general election campaigns found *zero* effect from persuasion tactics*

Testing is not...

- Polling/surveys
- Focus groups
- Advertising
- Media analysis
- Just “experimenting” with wacky, out-there ideas (rather, it can be a core part of regular operations)

When you hear “test” or “experiment” – think “comparison”

What sets a formal test/experiment apart from just collecting data or advertising is that you are consciously making a careful comparison.





An example test

For example: Motivating climate action—in Texas!



Oil Change International

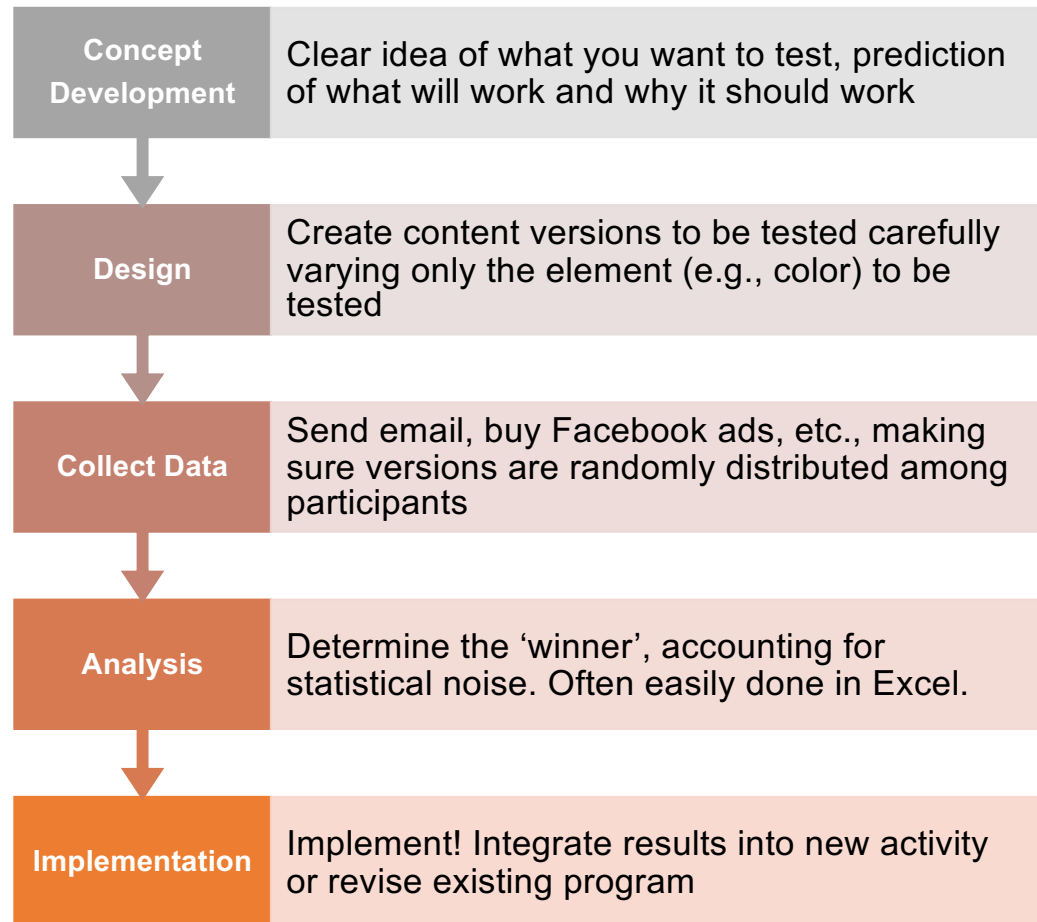
Sponsored • Paid for by Oil Change International

You're not alone. A growing number of Texans (now 71%!) agree we should do more to fight climate change, including a majority of Texas Republicans. However, U.S. oil and gas production is poised for an expansion that will be unparalleled

2 percent reduction in cost per engagement from added social norm language



Steps in the Testing Process

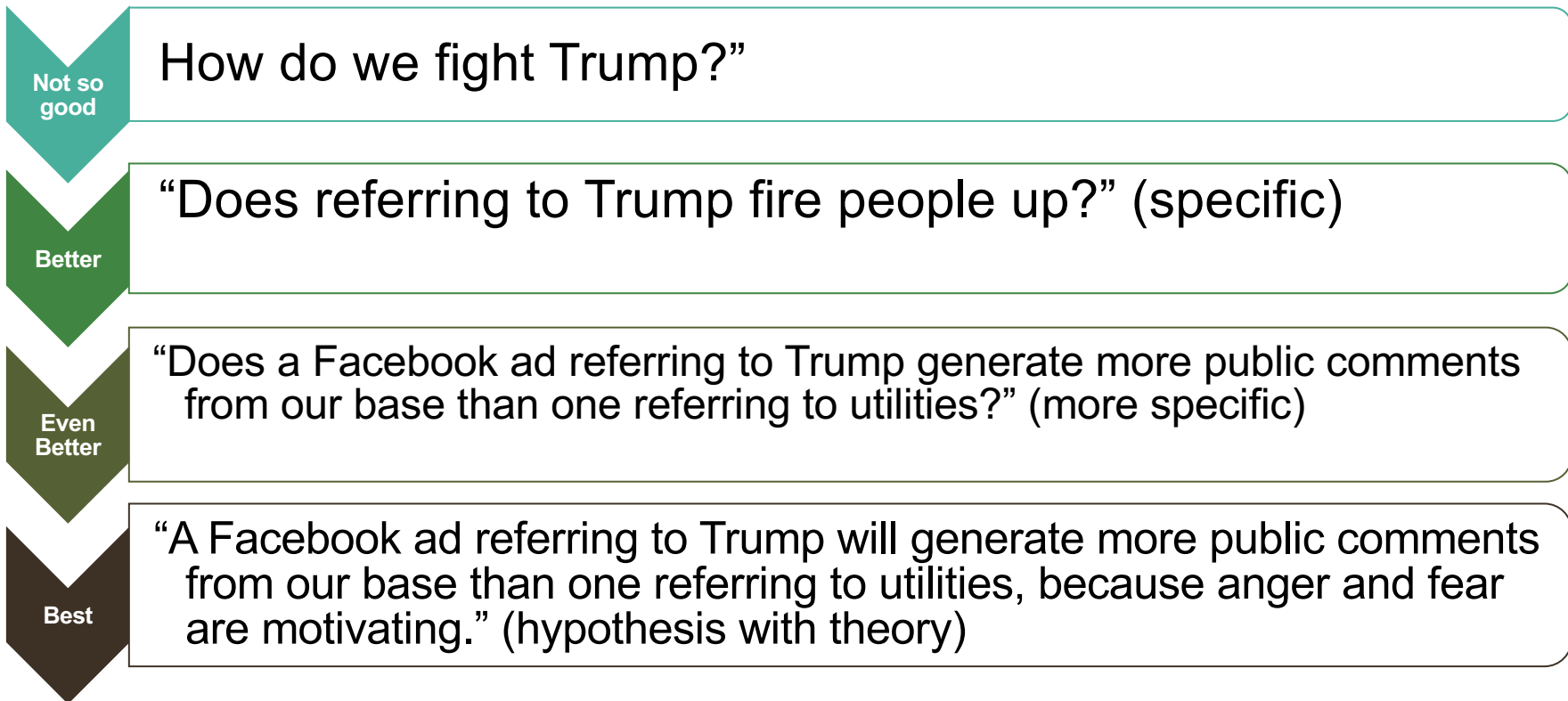


1. Concept Development

What is a question that, if you knew the answer *today*, would lead you to make decisions differently *tomorrow*?



Get from *general* to *specific* research questions



2. Design

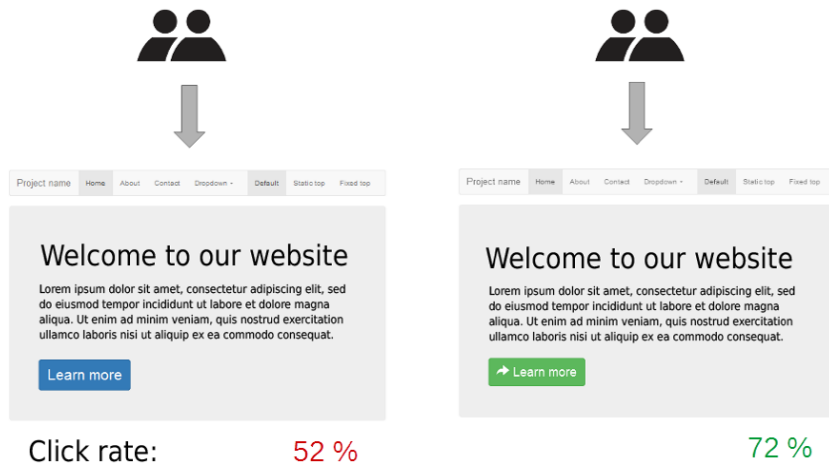


General testing design principles

- **Play with what you've got.**
 - Instead of investigating a totally new idea, test a variation to an existing program.
 - Ensures results will be actionable + allows you to compare results against an existing baseline ("we improved 25%").
- **Go big.**
 - Experiment with substantial changes, not small tweaks.
 - Tiny differences (in wording, etc.) rarely make a (significant) difference.
- **Focus on the relevant + actionable.**
 - How will your work *change* based on the results of your test?
 - What is a question that, if you knew the answer today, you would make decisions differently tomorrow?
- **Test on meaningful metrics, not vanity metrics.**
 - Beyond clicks and size, focus on impact.
 - How do you know your theory of change is working?

Two key aspects of rigorous testing design

Careful variation

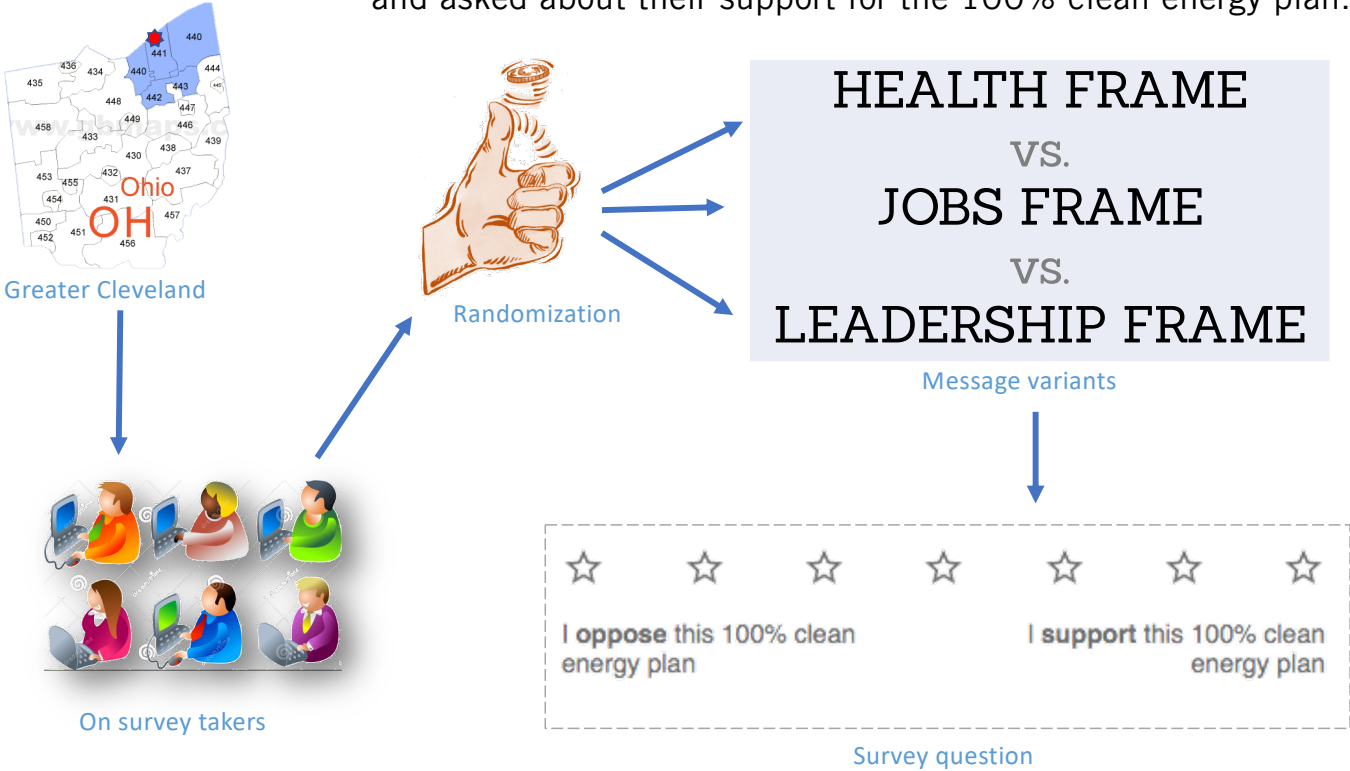


Random assignment

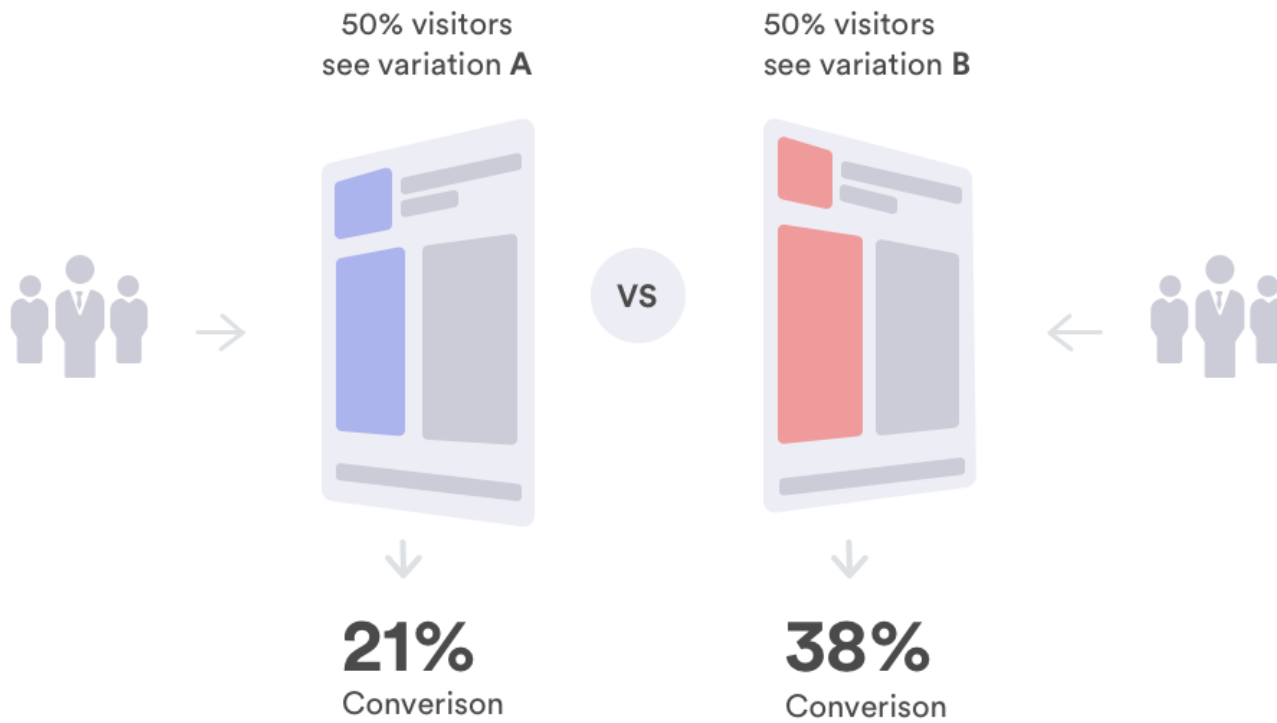


Example experiment design flowchart

Greater Cleveland participants each were randomly shown **one** message variation and asked about their support for the 100% clean energy plan.



A/B Testing



Credit: Orderhive

California EJ Alliance A/B Message Test

Fight Back Against Big Oil

Protect Community Health





CALIFORNIA
ENVIRONMENTAL
JUSTICE ALLIANCE

A/B Message Test

[A] Fight Back Against Big Oil

Environmental justice communities suffer from a range of pollution sources all at once. On top of extreme heat, poor air quality, or access to housing or jobs, all of these climate impacts are and will increasingly be felt in the same communities that have the fewest resources to adapt. **California's families, workers, and children are held hostage by the power of Big Oil and must be protected from dirty smog clogging their lungs everyday.**

[B] Protect Community Health

Environmental justice communities suffer from a range of pollution sources all at once. On top of extreme heat, poor air quality, or access to housing or jobs, all of these climate impacts are and will increasingly be felt in the same communities that have the fewest resources to adapt. **The health of California's families, workers, and children must be protected from dirty smog clogging their lungs everyday.**

3. Collect Data

What outcomes to test?

- **Vanity metrics:** Easy to measure, but not meaningfully linked to your strategy or theory of change
- **Meaningful metrics:** Those that inform decision-making towards accomplishing your goals

What do you want to know from your test?

- **What people think.** Use survey-experiments & respondent pools
- **What people do in the real world.** A/B test with email, Facebook, your site, your programs.
- **Both!** Conduct hybrid tests with a survey and a real world component

Which outcomes do you want to measure?

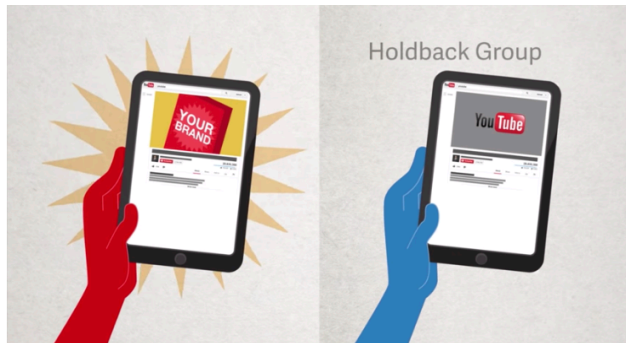


Changing Minds

- *Examples:* Persuading, informing, raising importance, counter arguing, reframing
- Harder to measure
- Methods
 - Placebo controlled tests
 - Pre-post measurement
 - ‘Lab’ pre-testing (e.g., Swayable, Google Surveys)
 - “Brand lift” surveys (within-platform)
 - Independent surveys of audience

Tools for testing persuasion

Brand lift studies & survey-based platforms



Google consumer surveys



SWAYABLE

**Which
outcomes do
you want to
measure?**



Online actions, engagement

- *Examples:* open rate, click rate, shares, likes, comments, traffic, sign ups, conversions
- Easy to measure within platforms, CRMs
- Methods
 - A/B tests
 - Dashboard stats
 - Tracking/conversion pixels
 - Referral link codes

Platforms for field testing

Measuring actions, not surveying to uncover attitudes

- Facebook
 - Popular testing platform, but design logistics can be tricky
- Email
- Online advertising: Google Adwords, etc.
- Testing of websites/landing pages via tools like Optimizely, Google Analytics, etc.
- Any systematic large scale campaign activity
 - Embedding tests in phonebanks/robocalls (e.g. patch-throughs), canvass scripts, volunteer engagement programs, etc.



4. Analysis

Analysis Options

DIY

- E.g., Excel, with Analysis Toolpak: <https://bit.ly/2HUXLz9>

Platform-based

- Tools such as Optimizely automatically analyze

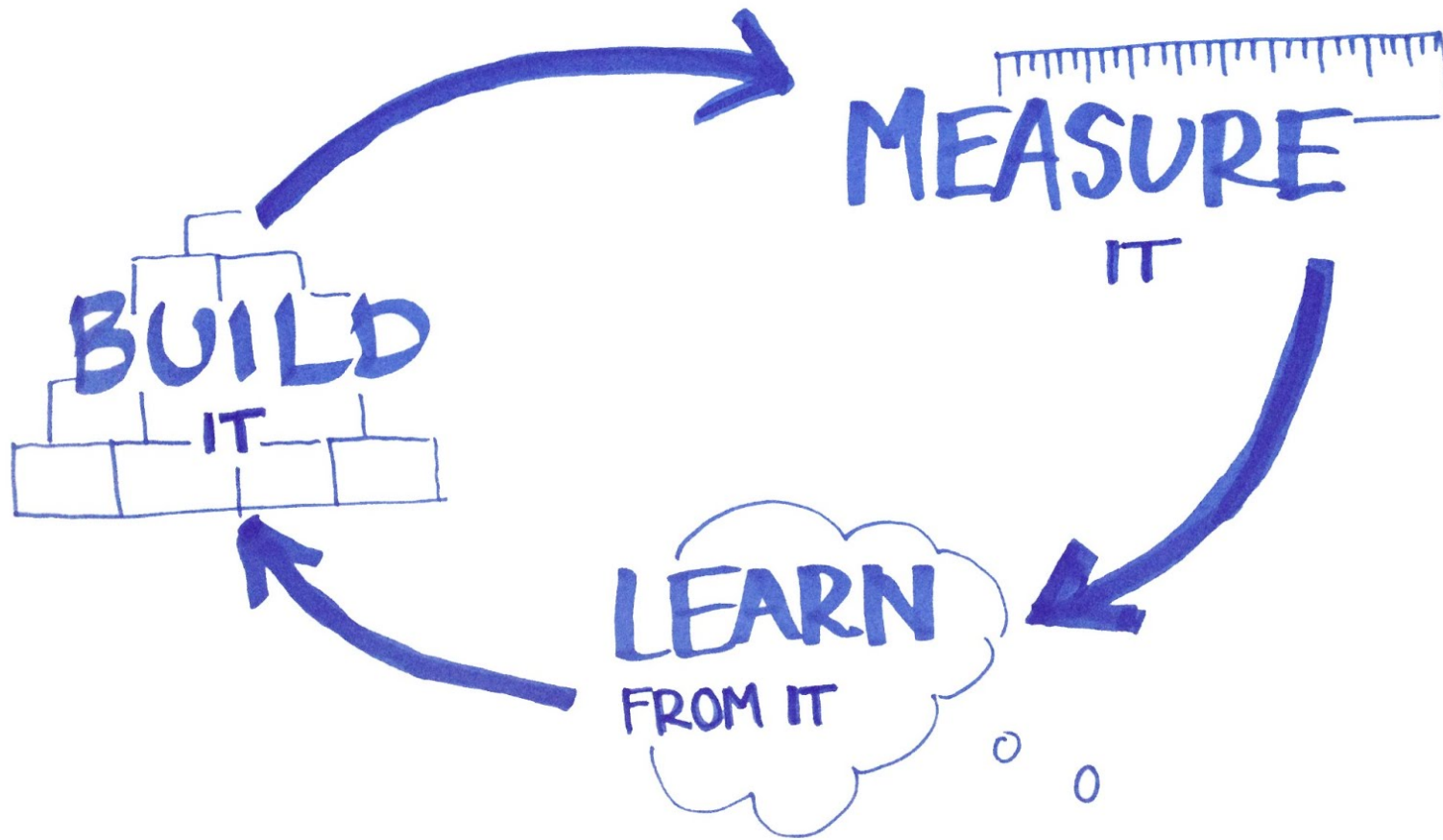
Outside help: The more sophisticated consultants can do this, including...

- Analyst Institute
- Collaborate with the Yale Program on Climate Change Communication's Facebook testers
- Your friendly neighborhood social scientist!

5. Implement!

Implementation tips

- **Buy in:** Involve everyone *early* who will need to sign-off and implement
- **Compare to a baseline:** Design experiment so it can show an X % improvement over how you already do things
- **Test grounded questions:** Before running test, ask yourself, “how would we do things differently if we knew the results?”
- **Be ready for failure, and improve:** Only 26% of average Optimizely tests worked (but 77% of those following best practices worked)
- **Promote!** Share test write-ups within your organization and with the Climate Advocacy Lab



Credit: Speckyboy