Experiment Cookbook

August 2019 | © Erik van der Pluijm | Wrkshp.tools

[1] Introduction

If you are an innovator or entrepreneur (and chances are that you are, since you're reading this), you know that building a new, successful product or business is both incredibly rewarding, and incredibly hard.

Achieving it means you'll not only need to learn about your customers and market, deeply understand their needs, and fulfill that need in an exciting way, but you'll also have to overcome … and even 'wiring problems' in the human brain, that conspire to make you take bad decisions.

For everyone involved in startups and new products, it makes sense to try to increase the success rate.

Unfortunately, nobody knows the exact recipe to build a successful business or product from scratch. There is no sure-fire step-by-step plan to follow, where if you check all the boxes, you'll end up with a success. It's not that easy.

However, there is one thing great innovators and entrepreneurs have known from the start, and that is to figure out what the biggest risks and roadblocks are, and then to try to solve those with small experiments.

Some of these innovators and entrepreneurs did this based on their own, private methodology, or even on a very well-developed gut feeling. That makes it hard to follow in their footsteps. Some of them used a more structured approach based on science.

That is exactly what Experiment Driven Innovation does.

The Lean Startup movement, started by Serial Entrepreneurs and Startup Gurus like Eric Ries, Steve Blank, Ash Maurya, and others, has become successful by taking the idea of the scientific method and using it to rigorously test business assumptions. This way of thinking helps you focus on the most important roadblocks first, and solve them in a way that reduces the risk.

Now, the idea of experimenting like this is straightforward. To most people it makes immediate sense. But when they try to run their first experiment, it turns out that it can be quite difficult to make it work and get clear outcomes.

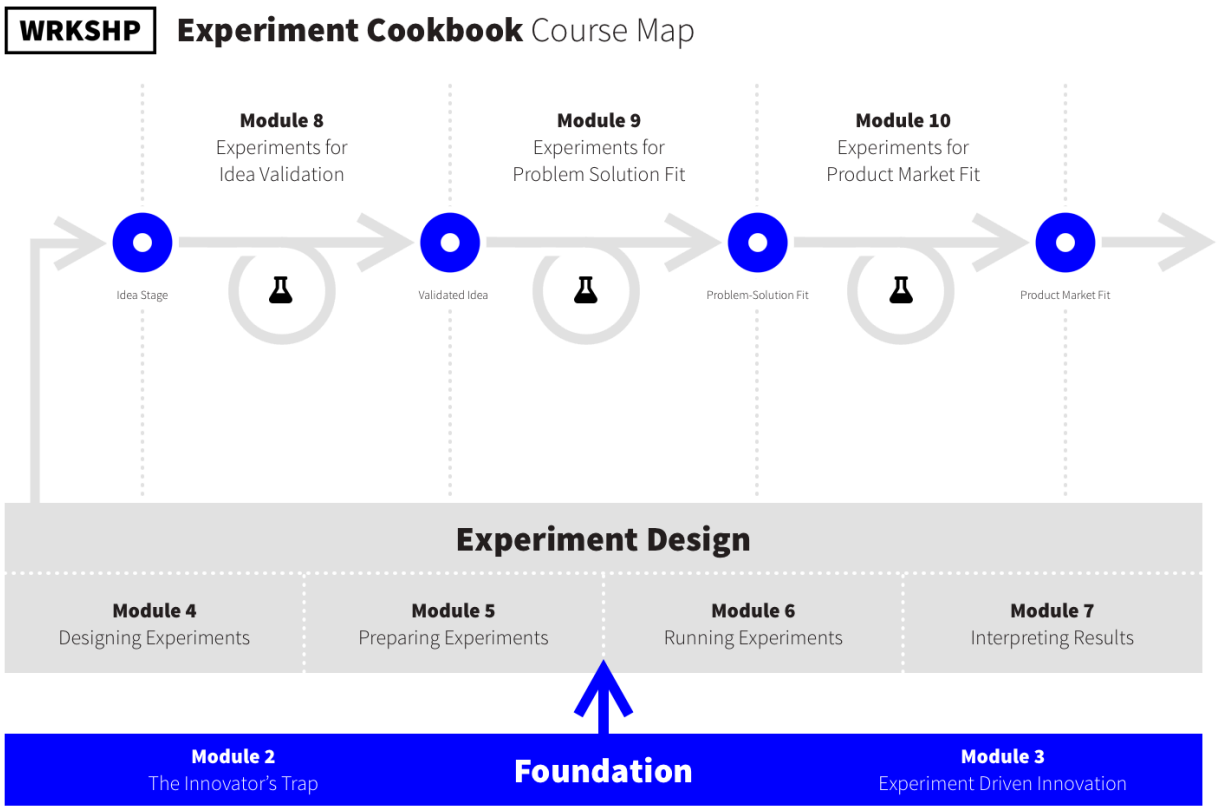
Most of the literature on Lean Startup thinking focuses on the theory, explaining why it is a good idea to experiment. Hardly any literature shows you the nitty-gritty details of running experiments in practice. What experiments should you run? How should you set them up? And when? What do the results mean? What is the best course of action? Tough questions to answer when reading the theory.

That means people are basing their potentially life-changing startup-decisions on flawed experiments.

When I started doing practical experiments with clients a few years ago, I constantly ran into these questions. What I really needed was an ‘Experiment Cookbook’. A list of recipes that I could draw from to build the right experiment for any situation.

This course aims to give you that cookbook. It will help you solve that practical problem and get you to running super effective experiments in no time.

[1.1] About this Course



This course comprises of four main parts:

1. An introduction in the theory and practice behind experiment driven innovation (Modules 1, 2, 3, 4)
2. A step by step guide to designing, running, and interpreting experiments (Modules 5, 6, 7, 8)
3. Guidelines per Innovation Stage (Modules 9, 10, 11) with Cookbook Recipes
4. Useful tools (Module 12)

Per Module

The structure of this course is as follows:

1. **Introduction** - General information about the course and structure.
2. **Why Experiments?** - Why Validation and Experiments are important. This module dives into background theory on the pitfalls of Innovation

3. **Lean Innovation** - What is lean innovation? What is Build-Measure-Learn? What are experiments?
4. **What Experiment Should You Run?** - Different types of experiments, how to choose depending on your assumption and where you are in the innovation journey
5. **Designing Experiments** - The nitty-gritty on designing experiments, riskiest assumptions, and hypotheses
6. **Preparing Experiments** - Guidelines for preparing your experiment
7. **Running Experiments** - Guidelines for running experiments
8. **Interpreting Experiment Results** - Guidelines for interpreting experiment data
9. **Experiments for Idea Validation / Problem-Market Fit** - How to adapt experiments for this stage plus recipes
10. **Experiments for Problem-Solution Fit** - How to adapt experiments for this stage plus recipes
11. **Experiments for Product-Market Fit** - How to adapt experiments for this stage plus recipes
12. **Tools** All Visual and Calculator Tools

This course is meant for innovators, entrepreneurs, founders, corporate intrapreneurs, consultants, and anyone else involved in innovation who wants to learn to run better experiments.

The course deals with qualitative and quantitative experiments, and offline as well as online experiments.

You can use the experiments in this course at any stage of the startup journey, but the examples focus on the earlier stages of idea validation and problem-market fit, problem solution fit, and product market fit.

The goal of the course is to:

- Give you a deeper understanding of experiment driven innovation
- Give you the tools, skills, and knowledge to design your own experiments
- Give you 25 practical experiment recipes you can use almost immediately
- Give you inspiration for your own experiment recipes

[1.2] What you'll learn

In this course, you'll learn:

- Background & Theory of Experiment Driven Innovation
- Choosing the best Experiment for your situation
- Designing Experiments
- Running Experiments
- Interpreting Experiments

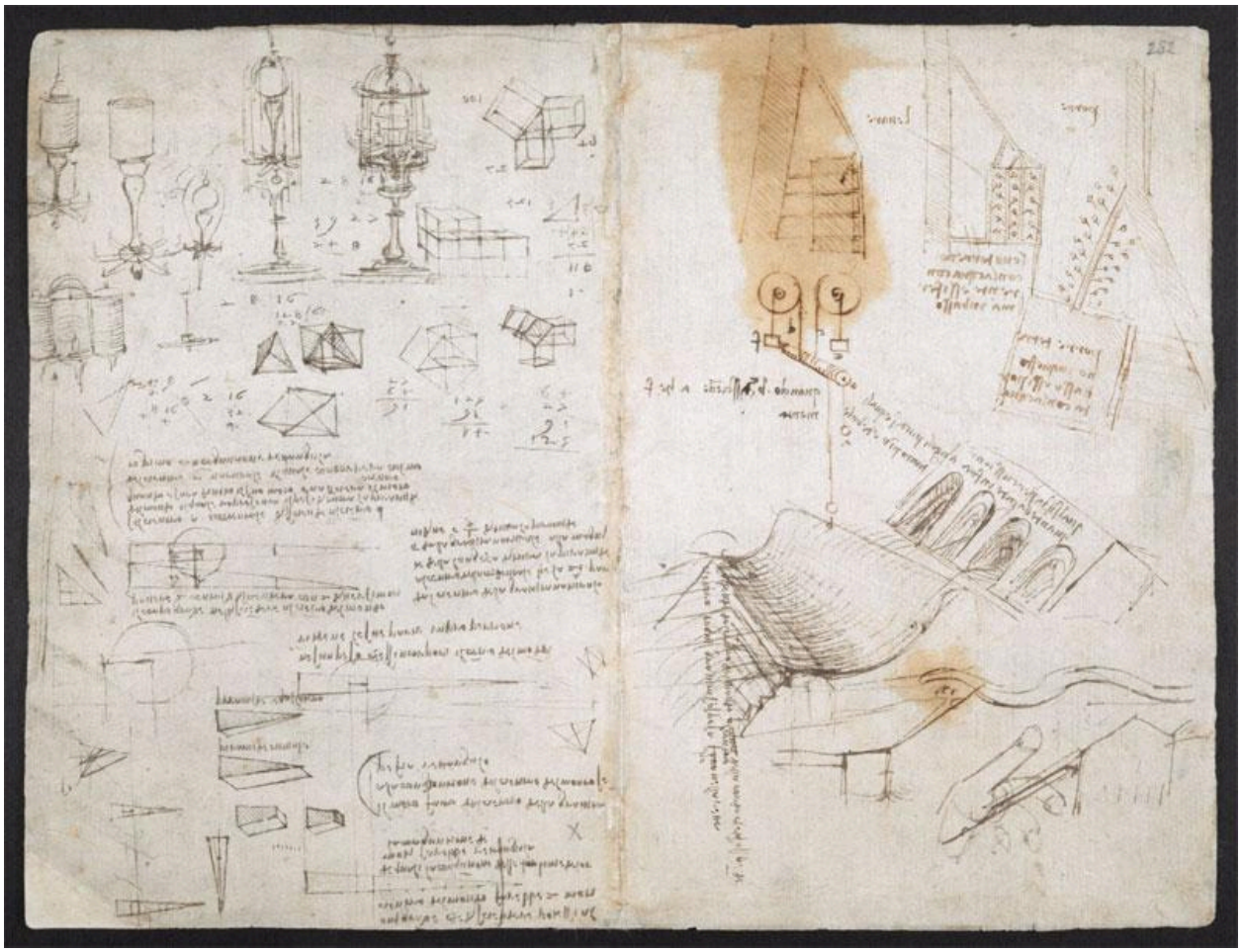
[1.3] How to get the most out of your course

This course will give you a lot of content to digest. The best way to integrate all of it is, fittingly, not by passively reading it, but by using it in practice. The best way to learn about experiments is to run experiments, and practice.

Be almost obsessive about it: see everything around you as an experiment, and even think about small daily actions using the terminology and concepts of experiments. In that way, you'll be able to make a shift in mindset towards a more objective, experiment driven approach.

Don't try to rush through the entire course at once. Stop and, well, experiment every now and then, and see if you really get it. I have added in exercises that you can use to practice, besides subjecting your immediate surroundings to (friendly) experiments.

[1.4] Your Logbook



Whenever you run experiments, if it is for this course, but especially if it is a ‘real’ experiment, please keep a detailed logbook. You will need it to reproduce the steps you took to get to your result. If you don't, learnings and ideas can turn out to be ephemeral. They will get lost in the noise of every day activity. Next thing you know, you'll be running in circles.

A logbook will save you when you have to base a difficult decision based on an experiment result.

Besides, it will help you when you want to build your own experiment recipes.

Give yourself a break. Keep a decent log 🙌

The following was adapted from Oxford University's Physics Undergraduate Practical teaching course handbooks by D.J.Jefferies, and underlines how strict scientific logbooks are kept. This is a good guideline for your own logbook, even if the practical side of business will mean that you'll probably be a bit less strict.

The importance of keeping a good experimental logbook.

Why is the experimental logbook important?

- It enables a complete reconstruction of the experiment or measurement at a later date. (It prevents 'experiment rot', where you're unable to trust the outcome of an experiment because you don't know how it was performed)
- It enables the work to be repeated for re-evaluation of the reported results.
- The steps that led to the success or failure of an experiment may be extracted.
- Investors or managers may ask for detailed reasoning that supports a tough decision based on experiment outcomes.

What kind of book for a logbook?

- A bound notebook with numbered pages is essential so that alterations, removals, and additions, are difficult to make at a later date. Alternatively, you can use something like evernote.
- The notes should be legible, in non-fading ink, and written as the measurement proceeds, not afterwards from memory of what happened.

Truth in reporting observations and experiments

Results should never be massaged to improve their fit with preconceptions. Negative results are as important as positive results. Facts should be recorded without comment, opinion should not feature at all. Uncertainties should be noted, unrepeatability recorded.

What should be recorded in the logbook?

In general, anything which allows someone else to repeat what YOU did in the experiment.

- Steps and procedures which are not obvious (The design and setup of the experiment)
- Precautions taken to minimise error
- References to other people's work, ideas, hints, and inputs.
- Parameters which might affect the outcome of the experiment.
- Equipment and tools used.
- Sketches of experimental setup.
- The date and time.
- The names of other people observing.

- Rough error analyses taken during the experiment, repeat observations of doubtful readings, calibration errors allowed for.

Data recordings

Logbooks record **data** as well as **methods**. The data should be ordered and logical. The instrument used should be recorded at every column of data readings, with an indication of the precision and the range setting of the instrument. Range changes should be noted.

Raw data should be recorded as well as processed data, so that algorithmic error can be spotted at a later date.

Charts, photographs, and printer output from the experimental apparatus can be cut and glued in the logbook wherever possible.

Validation of the logbook.

The logbook should be shown and results discussed with other team members, preferably on the same day that it is written.

[1.5] Checklist - Before you start

Before you really dive into the next module, consider the following checklist items:

✅ **Logbook** Start your experiment logbook
✅ **Learning Goals** Write down what you want to learn in your logbook

[2] Why Experiments?

The point of this chapter:

- Why do you need to use Validation
- Why is it important?
- Why does it work?

[2.1] The Innovator's Trap

You've probably heard a lot about validation, the Lean Startup, and running experiments already, which led you to sign up for this course.

But for those of you unfamiliar with the lean startup, or want a refresher, I'd like to dive into why validation is so important.

Zero to One (*)

It all starts with an act of creation.

Look around you, right now. Realize the following: once upon a time, everything you see around you did not exist. Computers, cell phones, chairs, clothes. Everything had to be created first. Think about it.

Such acts of creation are what is driving change in the world around us. They are the raw source of innovation. But what makes such an act of creation possible?

How does something go from a 'zero', nonexistent, 'impossible', crazy idea, to a 'one', something that is so obvious we take it for granted, affecting the daily lives of millions of people?

> Everything around you that you call life, was made up by people that were no smarter than you. And you can change it, you can influence it, you can build your own things that other people can use.
>
> — Steve Jobs

Video: Steve Jobs's vision of the world

*) 'Zero to One' is taken from Peter Thiel's book Zero to One - https://amzn.to/2TE8U7Q

The Reality Distortion Field

Making the transition from 'zero' to 'one' is hard. The people that attempt it, the innovators, are fighting the current. They must have complete belief in their ability to make the change happen. And through this belief, they create what has been called a '**Reality Distortion Field**'. In their world view, they are absolutely on the right track, and the change they are trying to make is completely obvious and natural. There is little place for self-doubt.

So far, so good. Believing in what you're doing is valuable. But there is a trap. It's very easy to turn a blind eye to any 'negative feedback' that is not in line with what you believe the world should be like.

The following action-reaction patterns may seem familiar:

- Customers say they don't like you product? Well, they just don't understand it yet!
- Sales targets are not being met? We just need more marketing!
- Investors don't believe in where you want to go? They're old school!

It's easy to have these 'knee-jerk' reactions to negative feedback.

The reason for this is that we humans have a tough time dealing with Cognitive Dissonance.

> In the field of psychology, **cognitive dissonance** is the mental discomfort (psychological stress) experienced by a person who holds two or more contradictory beliefs, ideas, or values. This discomfort is triggered by a situation in which a person's belief clashes with new evidence perceived by the person. When confronted with facts that contradict beliefs, ideals, and values, people will try to find a way to resolve the contradiction to reduce their discomfort
>
> [Wiki: https://en.wikipedia.org/wiki/Cognitive_dissonance]

It is very difficult to **both believe in what you're doing and be open to evidence from outside**.

When confronted with new evidence that conflicts with your belief system, it's tempting to reject the evidence rather than update your world view.

It is easy to respond to cognitive dissonance with dogma. Rather than believing in the **direction** you want to go in, you believe in the **particular solution** you have come up with. This means you're not only trying to change the world – **you want to do it in a very particular way**, and reject any evidence that does not support your views.

Many founder teams have fallen into this trap, dreaming up world-domination plan after world domination plan, spending thousands of hours or millions of dollars building something that people did not need, developing features that nobody cared for.

The world is a very complicated place, and '**betting the farm'** on thinking you know exactly how it works right from the start is very dangerous.

Examples:

Netflix

Justin Lokitz, one of my co-authors for the book Design A Better Business, describes a great example for a company everyone knows: Netflix.

> In 2011, the streaming entertainment giant, Netflix, decided to split its streaming and DVD business into two separate businesses, with separate names, and separate websites. These would be Netflix, the streaming service, and Qwikster, the DVD-by-mail service. On paper this idea probably looked great. By totally separating their subsequent business models, the company would be able to develop operational and marketing strategies specific to each. Makes sense.
>
> Actually, to the customer, it didn't make sense at all. Netflix delivers a set of services that are all about delighting customers. The very nature of Netflix's rise to prominence was its ability to continually address its customer's specific entertainment needs. The idea and resulting decision to split the company and services in two was never validated with Netflix's customers. Subsequently, shortly after the split, Reed Hastings, Netflix's CEO, made the following announcement, "It is clear that for many of our members two websites would make things more difficult, so we are going to keep Netflix as one place to go for streaming and DVDs."
>
> *(From: Design A Better Business)*

Better Place

Better place was a transportation company with the goal of completely revolutionizing the way we power vehicles by replacing petroleum powered engines with clean and affordable electric motors.
So far, that doesn't sound so different from Tesla. What is different, is that Better Place focused on a specific solution.

When Better Place launched in 2007, the loading time needed to give your car a good charge was huge, which was seen as a major problem for consumers, preventing them from switching

to electric cars. This was a major headache that needed to be solved if the company were to become successful.

Where Tesla works to improve the batteries themselves, improving load times and learning from customers with each car they sell, charismatic CEO Shai Agassi bet everything on a complex, capital intensive solution.

Better Place wanted to solve the problem with technology that would replace the entire battery compartment of a car in minutes in specially designed charging stations. To make this solution work a lot of moving parts were needed. Stations would need to be built, specialist technology created, and special swappable battery packs for cars built.

This approach led Better Place to burn around 850 million USD before going bankrupt in 2013, having delivered only 1500 vehicles and a handful of loading stations.

But even if they would have managed to pull it off, charging times have gone down drastically because of better batteries. The problem they focused on has gone (mostly) away, removing the base for their technology.

The long read on Fastcompany below tells a story where the reality distortion field became too strong, obliterating anything else.

Fast company: https://www.fastcompany.com/3028159/a-broken-place-better-place
Wikipedia: https://en.wikipedia.org/wiki/Better_Place_(company)

Webvan

Webvan was one of the very first online grocery delivery services. It was founded in 1996 and received around $600-800 Million in investments. Webvan had 3500 employees when it folded in 2001.

The most important reason that has been given for the failure is that the business model of Webvan had not been proven before they started aggressive expansion all over the US.

Today, shopping for groceries online seems so obvious it's hard to remember that in 1999 it was a different world. People were only getting used to the internet. Even Amazon wasn't the giant it is today. People were doing almost 100% of their shopping in actual stores.

So, even though the vision behind Webvan seems to have come true today, it could not yet exist (at scale) in the 90s. The mass customers that were needed to justify the scale of the operation were just not ready for it.

The small group of people that were ready for the service – the group called ‘early adopters’ in Geoffrey Moore’s excellent book Crossing the Chasm) – were not in Webvan’s crosshairs. They were looking further ahead in their spreadsheets and wanted to immediately focus on the mass audience. The business model they came up with would only work out when they would get a large audience to shop with them.

They ignored the fact that nobody knew at the time how to pull off successful e-commerce at scale, how to deal with logistical problems, and, the biggest unknown: how customers would react.

In their spreadsheet worldview it all worked out to huge profits.

Better validation and the ability to deal with cognitive dissonance and face the facts would have brought this to the surface and helped avoid this strategic disaster.

Conclusion: don’t bet the farm

The conclusion is to not bet the farm – at least, not all of it at the same time.

The examples of Webvan, Netflix, and Better Place show very large initiatives failing spectacularly. They make good stories, but the truth is that most failures never come to light: they are the small startups quietly failing before making a single mark on the world – let alone putting a dent in the universe.

Think of all the cases where smart people sat together, decided to start a company, and then stuck their collective heads in the sand and ignored any contradicting evidence.

Would you rather be right? Or successful?

It’s not roulette – it’s poker

Entrepreneurship is not about being ‘right’. It’s not about picking winners from the start. It’s not like **roulette**, where you put everything on double zero and hope for the best, with no plan B. That would be a poor strategy.

It’s much better to think about it in terms of **poker**. You have to play your hand in the best way you can. It is possible for a bad hand to improve, once you see more cards. And it is possible to play a bad hand and still win by understanding the game and your opponents.

> Life's too short to build something nobody wants.
>
> - Ash Maurya

Startups need to be ahead of the game, they're trying to find the next big thing. Their ideas are the opposite of currently accepted market needs.

That's why as a founder you need to confront reality on a daily basis. You need to become your own worst critic. Don't start drinking your own Kool-Aid. Don't skip steps of the innovation journey based only on the hubris-fueled belief that your idea is solid.

Instead, validate. Validate that **people experience and care about the problem** you want to solve. Validate that they not only want your solution, but they actually **need** it and **want to pay for it**. Validate that they **prefer your solution over all alternatives**.

And remember, you have proved nothing until you have your first 10 paying customers.

Examples:

AirBnB

AirBnB started in 2007 as 'AirBed & Breakfast', as an attempt by the founders to make some extra money while trying to make rent in San Francisco. They tried a number of things, including selling cereal boxes. Renting out air beds on their own apartment's floor was the thing that stuck, in the end–even while in the beginning it didn't really work out. They managed to balance their belief in the idea very well with different experiments, with different kinds of customers.

The AirBnB we know today was not something that was on the drawing board from the start. It grew from experiments. For instance, one successful experiment was to send a professional photographer to photograph airbnb's to make sure they looked great, which pulled in more users.

The AirBnB Story: https://amzn.to/31FncYH
Shorter version:
https://www.telegraph.co.uk/technology/news/9525267/Airbnb-The-story-behind-the-1.3bn-room-letting-website.html

Slack

In 2009, Stewart Butterfield, already founder of Flicker, had an idea for a new online game called Glitch with a new company called Tiny Speck. In 2011 they launched Glitch, a browser-based nonviolent MMO (Massively Multiplayer Online game) where players roamed the

minds of 11 cosmic giants. It bombed. The studio would be out of business. With just a month left, Stewart and his team decided to change direction and salvage what they could.

They took a good look at everything they had built in the past years. Was there anything left to convert into a viable product?

It turned out that while working on this game, they had been developing a chat tool to facilitate internal communication. When they hit their brick wall, that tool had been developed almost to the level of a finished product.

They launched it as ‘Slack’ and quickly found their first customers. In that way, they managed to turn around the company. (The funny thing is, that Stewart developed Flickr also as a pivot from a failed game… I wonder what his next game will be.)

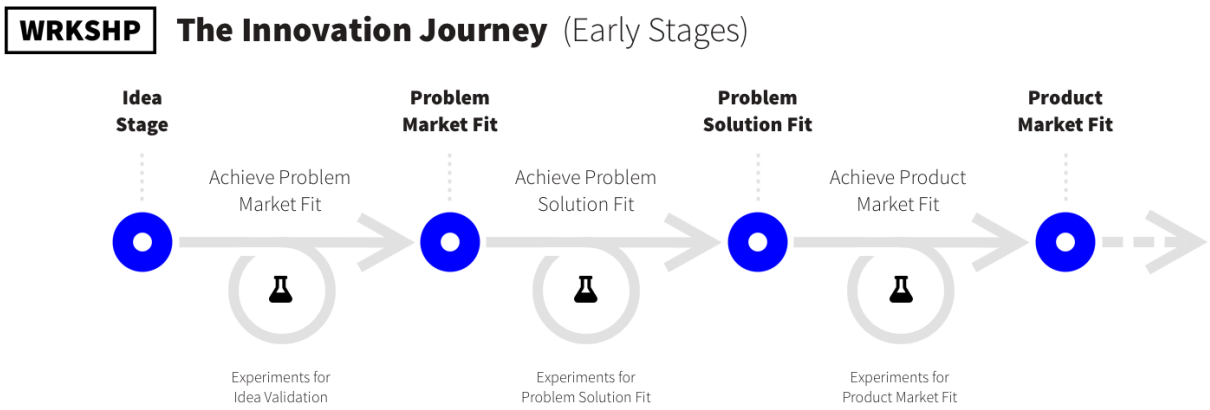
The Slack Origin Story on Techcrunch:
https://techcrunch.com/2019/05/30/the-slack-origin-story/?guccounter=1&guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_cs=RMofI-tHuWlL9pvTseR74g
On Fast Company:
https://www.fastcompany.com/3026418/this-story-about-slacks-founder-says-everything-you-need-to-know-about-him

[2.2] Exercise - Find Pivots

Think of three examples where companies or projects have changed direction, or tried something new. How did they do that? And how did it turn out? Did they change direction based on new evidence? What was needed to make them open to this change?

Think of one example where a company failed to take new evidence into account. What were the consequences? Did they fail to see it coming? Or did they ignore the new information?

[2.3] The Innovation Journey



This image shows the early stages of the innovation journey.

Traditionally, the innovation journey is divided into several stages. This course focuses on three three of the early stages of the journey, called Idea Validation or Problem Market Fit, Problem Solution Fit, and Product Market Fit.

In other literature these stages may be called slightly different, bet in essence it's always the same setup. Here's what the stages mean:

Problem Market Fit

From the start, beginning with an idea or first direction, the first stage is to find out if you can find people that experience and care about the problem. People that really need to have that problem solve. If you can prove this, you have achieved 'problem market fit': the problem you want to solve exists in the market.

(Problem Market Fit is described in more detail in module 9)

Problem Solution Fit

The next stage revolves around finding out which solution your customers prefer. What is the way in which their problem should be solved? Proving that you have a solution they prefer means achieving problem solution fit.

(Problem Solution Fit is described in more detail in module 10)

Product Market Fit

In the third stage, you will need to prove that your solution or product has a market, and you can find enough customers that really want your product.

The innovation journey continues with growing and scaling the business. Although you certainly can (and should!) use experiments there, the recipes in this course are targeted to these first three stages.

(Product Market Fit is described in more detail in module 11)

[2.4] Front load your Risks

In 1978, a small startup company was one of many to enter the new market for (personal) computer software. It was called Microsoft. When you think of Microsoft today, it's a giant company that has had a huge impact on the way we use computers. But in 1978, they had a handful of employees.

Microsoft employees December 7, 1978.

The question is, if you'd known the company back then, would you have invested? Probably not.

Why? Well, how would you have been able to tell that they were going to be hugely successful? At that point, they looked very similar to hundreds or thousands of other small startups working on computers, operating systems, and software. Most of those you have never heard of. Some of the ones that failed looked great at the start.

The lesson to take away is that most startups start out the same way, with a lot to prove.

If you look at the process of building a successful startup company, you can see that the cost goes up moving forward, while the risk to fail goes down.

At the beginning of your journey, your team is small. There are few commitments. The cost of failure is small. With a low investment of time and money and only a handful of people involved, the company is much more flexible and able to transfer to a more profitable approach.

Later in the journey, things are very different. You may have taken on funding, meaning your commitments are higher. You'll have a growing staff, and other recurring costs. Stakes are high. Changing course becomes like turning an oil tanker. If something goes wrong now, it really hurts.

Plotting the example for AirBnB from the previous module, it's clear their risk of failure has gone down drastically since 2007. When they first started out, they had basically the same odds as any other startup: a 90%+ failure rate within three years. But AirBnB managed to bring down their risk. Today, while it is possible that AirBnB would fail in the next three years, chances are very slim.

So, where does this initial risk to fail come from? Why do so many early stage startups fail? And what did AirBnB (or any other successful business) do to bring that risk down?

When AirBnB started out, they had no customers, no revenue, and definitely no proven business model. They only had an idea, with the assumption that it might work. They didn't know who their customers were. They didn't know if they could scale their model to other cities. They tried different things, some of which failed, until they found something that seemed to work. They **learned** about their customers, market, pricing, and all the other millions of aspects that underpin the success of AirBnB, **while they were building the business**.

And, crucially, they managed to apply what they learned to improve on their original idea, slowly changing it into something that is acceptable to massive audiences.

Because they started to learn from their customers from the start, they were able to focus on some of the biggest risks first:

- Would people be interested at all to switch from traditional hotel rooms to use AirBnB?

- What circumstances would be required to make them come back?
- How can you find people that are willing to rent out their place?

These questions were much more important in the beginning than immediately scaling up, building a super complex website, or getting international offices. Those things raise the cost of failure, without lowering the risk. Before scaling up, the risk of failure needs to go down first.

The smart thing to do is to **front-load any big risks involved**. Tackle those risks at the start of the journey, where being wrong has less impact and the cost of failure is low.

Why do startups fail?

CBInsights has made a study of the top 20 reasons startups fail. The biggest one (42%) is 'no market need'. That's right, 42% of the failed startups in their research failed (in part) because they failed to solve a problem for their customers. People simply did not need what they were selling.

CBInsights: The to 20 reasons startups fail:
https://www.cbinsights.com/research/startup-failure-reasons-top/

Among the top 20, 'pricing/cost issues' (18%), 'poor product' (17%) and 'ignore customers' (14%) are all major contributors to startup failure.

> It's very difficult to build a business around a 'nice to have' product, so you should keep your burn low while you iterate your core experience to make it a 'must have'.
>
> — Sean Ellis

It stands to reason that **avoiding these reasons for failure** is crucial if your startup is going to be a success.

What if it were possible to front load these risks and tackle them right at the beginning of the startup journey?

That is exactly what Lean Innovation does.

Serial entrepreneurs and Startup gurus like Steve Blank (Four Steps to the Epiphany), Eric Ries (The Lean Startup) and Ash Maurya (Running Lean) advocated running experiments and using data and objective metrics to kill the biggest risks to your fledgling startup right from the start.

The Four Steps to the Epiphany - Steve Blank - https://amzn.to/2KCi6Xo
The Lean Startup - Eric Ries - https://amzn.to/2Hc2zeN
Running Lean - Ash Maurya - https://amzn.to/31FgaDh

The theory sounds great. But, if you're like me, in practice, it turns out to be difficult.

What experiments should you run? How should you set them up? And when? What do the results mean? What is the best course of action? Tough questions to answer when reading the theory.

When I first started doing practical experiments with clients some years ago, I constantly ran into these questions. What I really needed was an 'Experiment Cookbook'. A list of recipes that I could draw from to build the right experiment for any situation.

This course aims to give you that cookbook. It will help you solve that practical problem and get you to running super effective experiments in no time.

But before we dive in and start designing experiments, it's vital to learn a bit more about how to prevent falling in the Innovator's Trap. After all, even if you're armed with the best experiment recipes, if you start drinking your own Kool Aid you'll still make bad decisions.

[2.5] Avoiding the Innovator's Trap

If front loading the biggest risks is so logical, how come we still see failures like Better Place?

Writing off teams and founders who fall into this trap as '*stupid'* would be doing them a grave injustice. Sure, inexperience and bad decisions certainly play a role, but I believe anyone can fall into the innovator's trap.

Why do smart people fall into the Innovator's Trap?

Because they are **wired** that way.

There are certain inherent flaws in human thinking that push us towards making bad decisions.

Relying only on a strong Reality Distortion Field is one of these **'wiring problems',** that can seriously hamper your ability to be successful with your startup. But there are more:

1 Loss Aversion

In cognitive psychology and decision theory, loss aversion refers to people's tendency to prefer avoiding losses to acquiring equivalent gains: it is better to not lose $5 than to find $5. The principle is very prominent in the domain of economics. (source: Wikipedia)

We humans don't like losing. In fact, we hate it. We hate it so much, that we rather spend time and energy preventing or avoiding a loss than trying to create extra gains. This is something you can see happening in any large company, but it also happens to founders, when they spend more time preventing potential losses than looking for new revenue streams. Don't spend a day removing a $10 cost — use the Pareto principle and go for the biggest ones.

Loss aversion can make you play it safe, stay inactive. It also may make you defer decisions to later. This inactivity is paralyzing for startups.

When getting negative feedback may have implications for your job security, it may seem way better to work on your idea some more, polish it a bit, spend more time, than to risk getting bad feedback. At least, that way, when you do get bad feedback, you can say you have done everything you can to build the best possible product, right? Just look at how much work went into it! This response to uncertainty is what I call **'planoholism'**, being addicted to planning and reports, overthinking everything.

(Read more about planoholism:
https://medium.com/wrkshp/stop-being-a-planoholic-9c70a7d34f06)

> Asking feedback right away is like pulling off a band aid. The anticipation is way worse than the actual thing.

A great remedy is to force yourself to ask feedback right away, incorporating it in your process. Asking feedback right away is like pulling off a band aid. The anticipation is way worse than the actual thing. The sooner you get it over with, the sooner you can take action. The sooner you know your customers don't like your idea, the sooner you can come up with a better one. Staying in your ivory tower, pretending everything is A-ok, will definitely not help you.

2 Ego

Another reason people don't want to get feedback is to avoid hurting their ego. If they have built their self esteem and status on being 'right', they will need their **original idea** to succeed. It drives them to champion a particular solution, rather than falling in love with the problem.

Everyone attacking that idea, any negative influence, must go. They'd rather be 'right' than successful - or rather, they'd rather run the risk of being unsuccessful than listening to feedback and changing their idea.

Everyone can fall victim to this effect. We all have egos. It's important to recognize when your ego gets in the way.

3 The Dunning Kruger Effect

> [T]he **Dunning–Kruger effect** is a cognitive bias in which people mistakenly assess their cognitive ability as greater than it is. It is related to the cognitive bias of illusory superiority and comes from the inability of people to recognize their lack of ability. Without the self-awareness of metacognition, people cannot objectively evaluate their competence or incompetence
>
> [Wiki] https://en.wikipedia.org/wiki/Dunning%E2%80%93Kruger_effect

This wiring problem works both ways. Inexperienced entrepreneurs may fall into the trap that they think they know what they're doing, with a false sense of security, while experienced entrepreneurs may be overly insecure and too careful. The point is that it is very difficult to evaluate for yourself how competent you are. The only good way to do that is to get feedback.

4 Not Invented Here (NIH)

A variation on the Ego theme, the Not Invented Here Syndrome pops up when people (or teams) only believe in the value of their own ideas and work, failing to build on what others have done. This can happen because they are more familiar with their own work, creating a false sense of superiority, ignoring outside trends and opportunities.

For startups, especially in the early stage, it makes sense to force yourself to the opposite point of view:'PFE', or 'Proudly Found Elsewhere'.

5 Group Think

> Groupthink is a psychological phenomenon that occurs within a group of people in which the desire for harmony or conformity in the group results in an irrational or dysfunctional decision-making outcome. Group members try to minimize conflict and reach a consensus decision without critical evaluation of alternative viewpoints by actively suppressing dissenting viewpoints, and by isolating themselves from outside influences.
>
> [Wiki] https://en.wikipedia.org/wiki/Groupthink

Startups and innovation teams can be quite vulnerable to groupthink, because they operate in an environment of uncertainty and cognitive dissonance. It can be very tempting to find support within the group, and adopt the same viewpoint.

One way of countering this is to appoint a 'devil's advocate' whose role it is to challenge and question the direction you're moving in. Rigorously using experiments is another way of doing this, since they bring in an outside perspective through data.

6 Single Right Solution Thinking

Single Right Solution Thinking is a paralyzing condition that is all too common. It's perfectly natural for a team to develop tunnel vision, especially when you're working hard and are constantly under the gun, trying to make ends meet. There also is a natural tendency to follow habits, and to do things in the same way they have always been done. That is not necessarily bad in itself, just make sure you are not missing some awesome opportunities.

As an entrepreneur, it's your job to be able to recognize tunnel vision, take a step back, and see if there is another, easier, simpler solution. Dare to question if you're on the right track. Whenever there is a situation where it seems there is only one 'right' way to progress, make sure you check if it is really true that you have only one option.

> Creativity involves breaking out of established patterns in order to look at things in a different way.
>
> — Edward de Bono

7 Availability Heuristic

> The availability heuristic is a mental shortcut that relies on immediate examples that come to a given person's mind when evaluating a specific topic, concept, method or decision. (source: Wikipedia)

This mental shortcut will influence your decisions by making you only consider immediate examples. This means, that if you're drawing examples only from a narrow experience, you may make bad decisions. Don't assume that because three of your friends experienced a certain problem, everyone has that problem. Actively try to find counterexamples and do research.

The Availability Heuristic shows up when you have a lack of diversity in your team. Diversity in backgrounds, worldviews, and ways of thinking makes it harder to believe that everyone in the world thinks just like you do. If you're the statistically average startup group of white, highly educated young males, this can quickly become a big problem. The remedy is to actively look for viewpoints that are different.

8 Sunk Cost Fallacy

The Sunk Cost Fallacy makes it harder to change course after you have invested heavily in a course of action. It pushes you to throw good money after bad money: because you have invested thousands (or millions) on building an underperforming project, it is tempting to 'recover the loss' and throw more money at it to fix it. In many cases, this is a bad idea, and it's better to cut your losses.

When you find out you've been spending on a losing idea, try to calculate what it will cost you to continue vs what it will cost you to stop — and what opportunities you'll **gain** by changing course. Don't be throwing good money after bad.

9 Confirmation Bias

> Confirmation bias, … , is the tendency to search for, interpret, favour, and recall information in a way that confirms one's preexisting beliefs or hypotheses. (source: Wikipedia)

If you're an early stage startup, confirmation bias might be the most important cognitive bias to learn about. When you're starting out and are trying to validate your idea and find problem-solution fit, confirmation bias can severely hurt you. Make sure you keep trying to expressly look for signals that do not agree with your beliefs and hypotheses. If you can't find any customers that hate your solution, then you're on to something.

For anyone getting involved with experiments, Confirmation Bias is perhaps the most important wiring problem to be aware of. When running experiments, always look out for the tendency to reduce your cognitive dissonance and (selectively) find facts that support your assumptions. Force yourself to look for the facts that do not support them.

10 Cognitive Biases that can Kill your Startup:
https://medium.com/wrkshp/10-cognitive-biases-that-can-kill-your-startup-270a8c70e763

> "If You're Not Embarrassed By The First Version Of Your Product, You've Launched Too Late"
>
> -- Reid Hoffman (Founder of LinkedIn)

It is vital to fight these built-in preferences every step of the way. You'll need to rewire yourself and your team and inoculate yourself to the innovator's trap. One way to make that easier is by using Experiment Driven Innovation. This method can go a long way to prevent wiring errors by enforcing looking objectively at outside data.

[2.6] Exercise: Wiring Problems

For each of the ‘Wiring Problems’, come up with one occasion where you or team members have fallen victim to it. How did it happen? What might you have done differently?

- Reality Distortion Field
- Loss Aversion
- Ego
- Dunning-Kruger Effect
- Not Invented Here
- Group Think
- Single Right Solution Thinking
- Availability Heuristic
- Sunk Cost Fallacy

What other ‘wiring problems’ can you think of that hold you and your team back from going out there and testing your ideas?

[2.7] Reading List

Books

- Zero to One - Peter Thiel
- The Hard Thing about Hard Things - Ben Horowitz
- Crossing the Chasm - Geoffrey Moore
- The Four Steps to the Epiphany - Steve Blank
- The Lean Startup - Eric Ries
- Running Lean - Ash Maurya
- The AirBnB Story - Leigh Gallagher
- The Upside of Irrationality - Dan Ariely
- Thinking Fast and Slow - Daniel Kahneman
- The Art of Thinking Clearly - Rolf Dobelli

Blogs

- Stop being a Planoholic
- CBInsights: The to 20 reasons startups fail
- The Slack Origin Story on Techcrunch
- The Slack Origin Story on Fast Company
- Fuji's Ultimate Pivot
- How Kodak Failed
- The Story behind AirBnB
- Successful Startup Pivots
- 10 Cognitive Biases that can Kill your Startup

[3] Lean Innovation

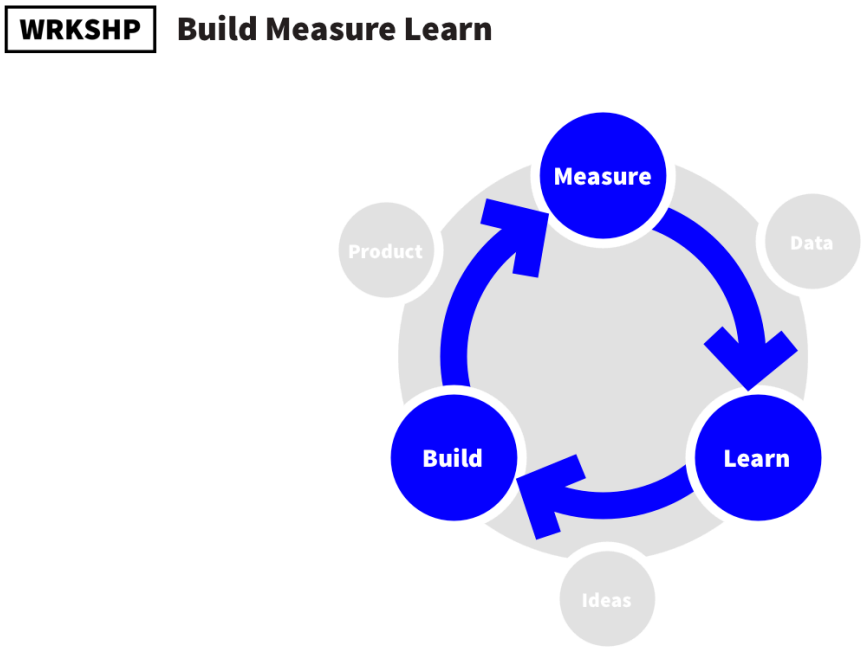
[3.1] What is Lean Innovation?

Some of the biggest risks an early stage startup faces revolve around finding, understanding, and then fulfilling a real customer need, it is very important to tackle these risks early on. You must know for sure that you have found a problem your customers care about and need to have solved, before moving on to build anything substantial.

You need to **learn** as much as you can. But how do you learn?

> The fundamental activity of a startup is to turn ideas into products, measure how customers respond, and then learn whether to pivot or persevere. All successful startup processes should be geared to accelerate that feedback loop.
>
> - The Lean Startup

[3.2] How do you learn? The Build-Measure-Learn Loop



A **'feedback loop',** mentioned in the quote above, is the quintessential part of any learning system. Learning happens when you are able to analyze feedback from the results of an action, and are able to adjust your next action to make use of that feedback so that you can reach a goal.

This is true for toddlers, adults, computers, and organizations. If there is a feedback loop, the person, or system, can be said to learn.

It will improve its performance over time, by learning from the results of its actions.

The Build-Measure-Learn loop is a popular way to describe a learning loop for organizations.

It has three steps:

- Build
- Measure
- Learn

Build

In the ‘build’ step, you perform an action that has an effect on the world outside. This could mean doing something like creating a first version of a product, and launching it.

Measure

Next, in the ‘measure’ step, you observe what happens. Are people using the product? Do they buy it?

Learn

Finally, in the ‘learn’ step, you analyze the results from the measure step. Was what you observed in line with your expectations? What can you change? What assumptions about your customers and product do you need to change, now that you found this new evidence?

Iterate

Then, the loop starts again. You modify the product, and launch a new version. Again, measure what happens, and then learn from your results. Keep going through this loop until there are no clear improvements to be had.

Even without any further formalisations of how you use this loop, using it will help you improve your startup’s success rate. Making the process more explicit makes it even more effective.

The basic build-measure-learn process does not tell you what you should be looking for in the measure phase, after you complete the build phase. It doesn’t tell you how to interpret the results and learn from them. And it doesn’t help you define what changes to make in the next build phase.

Experiments

This is where validation experiments come in. (Loosely) based on the scientific method, they help you focus on specific assumptions you need to test, give you ways to test those assumptions, and provide a clear outcome whether your assumption is correct or not.

Without a more formal method, using build-measure-learn is a bit like stumbling around blindfolded, randomly thrashing about trying to hit a piñata.

Using build-measure-learn with experiments makes it much more effective. Rather than using the ‘build’ stage to build actual working products, it is used to come up with a well-defined experiment to test a risky assumption. To test the assumption, usually something needs to be

created. This can be any prototype from a list of questions or a one-page flyer to a full fledged MVP, depending on the situation. Then, in the measure phase, the experiment definition guides exactly what to measure and how to record the results. Finally, in the learn phase, the data is interpreted according to the experiment setup, and a conclusion about the risky assumption can be drawn.

The fact that the experiment is well-defined helps inoculate you against wiring problems (such as confirmation bias) that inevitably show up, helping to make decisions with confidence.

Lean Startup

By combining validation experiments and the build-measure-learn loop, the lean startup defines a clear, actionable method that will bring you from uncertainty and high risk of failure toward certainty, and helps you focus on the most important assumptions, so that you can front load your risk.

[3.3] Learn by validating your assumptions

The Build-Measure-Learn loop describes the entire lean validation process. You have an assumption, based on that assumption you build something, put it out there in the real world with customers, measure their feedback, and learn from it to see if your assumption was correct. Then you do it again.

In short, validation means challenging your assumptions.

In the startup journey there are many things you can validate. There are two broad categories: you can validate either 'internally' or 'externally'.

Internal Validation

Internal validation is the least time-consuming type of validation, involving research and pen and paper. Think about market research, calculating business cases, and looking at benchmark numbers for conversion rates, prices, and so on in your market.

***Warning**: because internal validation involves using your own ideas and assumptions, you run the risk of buying into your own story. This can lead to confirmation bias. Try to be critical and check your findings with others.*

External Validation

External validation is more involved than internal validation, but it is also way more rewarding. To do external validation means to talk to (or interact with) (potential) customers. This is where you can really challenge your own assumptions and biases.

Watch out for confirmation bias

People tend to interpret ‘validation’ as simply a way of getting feedback to see if they are on the right track. However, while getting feedback and measuring is absolutely the best way to move forward, there is a warning attached. Don’t try to **validate** your ideas: you’ll run the risk of seeking confirmation bias, another ‘wiring problem’ entrepreneurs face every day. Try to **invalidate** your assumptions instead!

If you’re not actively trying to find evidence that invalidates your assumptions, you run the risk of missing that evidence. You'll only be looking for confirmation. And that will hurt you later on.

[3.4] Experiments: Structured Validation

Bring on the ‘science’!

Now that the Build-Measure-Learn loop and Validation have been discussed, it’s time to think about **how** to validate.

To learn about the world around you, you need to know what part you understand and what part you don’t understand.

The easiest way to do this is by taking an assumption you have about the world and seeing if you are correct.

To use the story of Columbus sailing west to reach India as an example, Columbus’ assumption was that the world was round. Based on his assumption, Columbus made a prediction: that it was possible to reach India by sailing west. Presumably, most people at the time had a very different assumption which led them to predict sailing west would mean falling off the edge of the world. Columbus then tested this assumption by sailing west.

If I assume people like my new product, I can find out by showing it to them. I can then measure how they respond. If my assumption is correct, I would expect a large majority of the people to say they like the product.

As can be seen from both examples, to test an assumption, it is necessary to make a **prediction**.

If the result was in line with that prediction I can be confident (to some degree) that the assumption was correct.

This prediction is what we will call the **hypothesis** of the experiment. The data you get out of your experiment will allow you to either confirm or reject that hypothesis, and thus validate or invalidate your assumption.

The Point of all this Structure

Perhaps you have the feeling that all this structure stinks of bureaucracy, and that it slows you down. All that time defining hypotheses while you could have been out there running experiments. But a minimum of structure is vital to your end result.

The main reason to define a ‘formal’ structure for your experiment is to avoid the ‘wiring problems’ discussed in Module 2, specifically Confirmation Bias.

Without a clear structure, it is very easy to run an experiment and selectively interpret the results to match with what you assumed to be true. You're falling into the confirmation bias trap.

It's a bit like playing pool without calling pockets. In pool, potting balls can be quite easy. But a true test of skill, as opposed to dumb luck, is to call the pocket you're going to put the ball into before you hit it.

The only good way to avoid confirmation bias in experiments is by doing something similar.

Define a clear hypothesis **before** running your experiment, so that it becomes impossible to 'fudge the numbers' afterward.

Logbook

Be 'scientific' about your experiments. Follow a strict protocol, so that each experiment is performed in the same way and you are able to compare the results. It also means, you'll need to take records. That's why we have been asking you to keep a **logbook**!

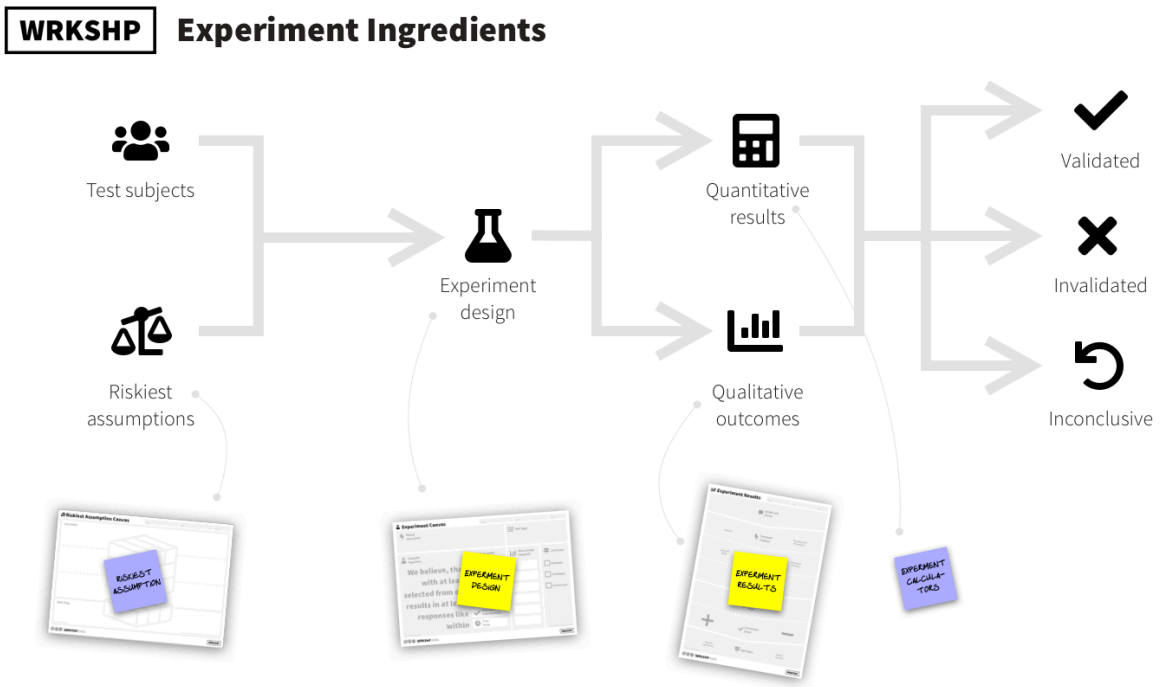
Especially in the early stages, when your experiments are still qualitative, it can be very difficult to interpret the results correctly. Not writing down the results and conclusions will only make that worse.

Not keeping a good log will only mean you'll move around in circles, having the same discussions over and over again as you start to doubt the results of past experiments.

Being able to look back and check your log for the conclusions, and what they were based on, will save you tons of time.

[3.3] The Ingredients of an Experiment

Every experiment shares a number of basic ingredients.



Every validation experiment contains the same ‘ingredients’, shown in the graphic above. Below, each of these ingredients is described in more detail. In the module ‘designing experiments’, each ingredient is tackled in detail.

The ingredients are:

- Riskiest Assumptions
- Falsifiable Hypothesis
- Method
- Prototype
- Protocol
- Data
- Conclusion

Riskiest Assumption
The Riskiest Assumption is the thing you want to validate. It is important to select the right assumptions to test first, and luckily, there are some rules of thumb you can use for each stage of your innovation journey. This will be explored in depth later in the course.

Falsifiable Hypothesis
The Falsifiable Hypothesis is a statement that is derived from the risky assumption and defines exactly what you are going to test, how you will do it, with how many people, and when the

result will be positive or negative. It is very important to quantify your hypothesis, and specify in detail how many respondents you need, how long it will take, and what you will count as a positive result.

Method
The Method is how you run the experiment. How do you set it up, how do you select respondents, where do you run the experiment, etc.

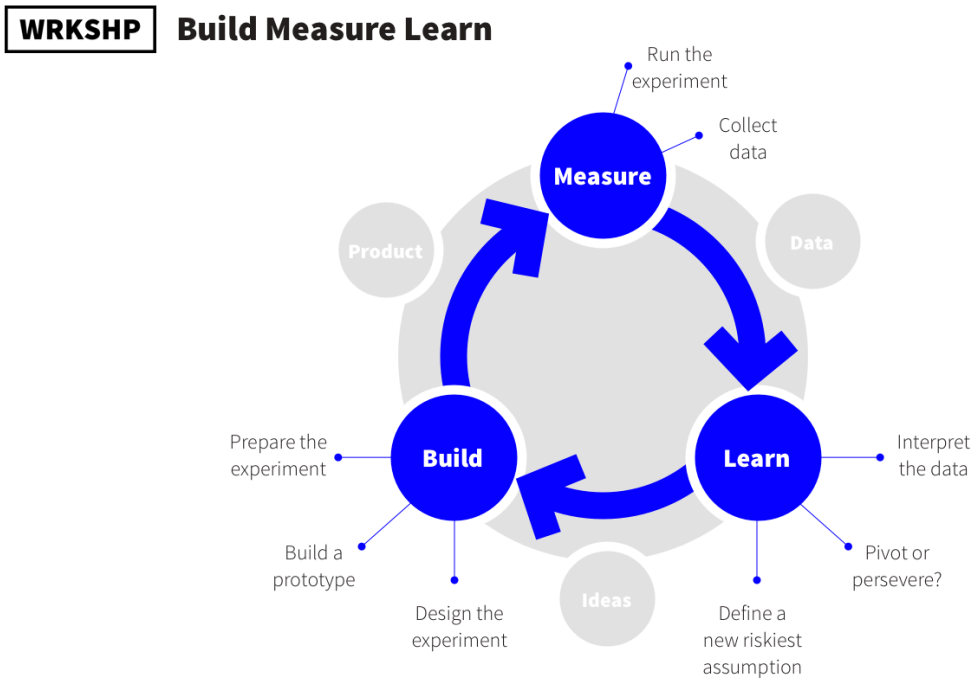
Prototype
Usually, when running an experiment, you will need some kind of artefact to test your hypothesis with. It could be as simple as a list of questions you discuss with respondents, or as complicated as a minimum viable product. This artefact is what I'll call the 'prototype'. There are different types of prototypes that suit different types of experiments.

Protocol
The Protocol defines how you derive a score from the results you get back. Which results count? When is a result positive? It can be simple (count all page views vs conversions on a landing page) or more complex (count all cases where people are 'positive' about product X - which means you'll need to define what 'positive' means)

Data
Data is what results from the experiment. The raw responses, page views, clicks, or other measurements you have taken during the experiment. Recording your data properly and storing it somewhere safe are super important.

Conclusion
The Conclusion is the decision you take after the experiment has run its course. Depending on how the actual data stacks up against the hypothesis, you'll either conclude the assumption is **validated** or **invalidated**. When the data is insufficient, the experiment could also be declared **inconclusive**. Based on this conclusion, you'll either **persevere** (continue on your current course), **pivot** (go back to the drawing board), or **redo** the experiment or find another way to test your assumption when the results are inconclusive.

[3.4] Experiments and Build Measure Learn



Now that the types and ingredients of experiments are a bit more clear, it is time to look at how they fit together with the Build Measure Learn cycle. How do you string experiments together to form a learning journey, leading your startup to product-market fit?

There are a few extra pieces to add to the Build Measure Learn loop.

Risky Assumptions

Your collection of risky assumptions is a reflection of what you know about the world, and what you need to prove to be true if your idea is going to work. During your learning journey, you'll regularly need to go back to your risky assumptions and look at what is the biggest risk right now.

Experiment Design

Designing new experiments falls in the build phase. When you design your experiment you'll usually also need some kind of prototype or artefact that you can use in the experiment. In the build phase, that is also created. The build phase should be as short as possible. What is the fastest way to get to the measurement stage?

Running Experiments

The measure stage is where you run experiments, gather data, and store results.

Interpreting Results

In the learn step, the measurements are interpreted. This is where you first of all decide if the result you found is usable: can you make a confident decision? Is the result significant? When you know that, you can start drawing conclusions.

Pivot or Persevere

This is where you apply what you have learned to decide what your next step will be. Do you continue on the same track, or do you need to pivot and take a different direction? This influences what your next riskiest assumption will be, and how to test it with an experiment.

[3.5] Reading

Books

- The Four Steps to the Epiphany - Steve Blank
- The Lean Startup - Eric Ries
- Running Lean - Ash Maurya
- The Startup Owner's Manual - Steve Blank

[4] Selecting Experiments

[4.1] Types of Experiments

There are two major groups of validation experiments depending on what you are trying to achieve.

- Qualitative / Exploration Experiments
- Quantitative / Validation Experiments

The distinction is made based on what you are trying to validate. Are you trying to validate a ‘known unknown’? Or an ‘unknown unknown’?

> There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns — the ones we don't know we don't know.
>
> -- Donald Rumsfeld

Each experiment recipe in the Experiment Cookbook is either qualitative or quantitative (or both).

Unknown Unknowns: Qualitative / Exploration Experiments

Especially in the beginning of the innovation journey, when you’re still exploring what your customers response to the problem is and have a lot of assumptions, but very few facts, you should be looking for ‘unknown unknowns’.

In that stage, thinking you know anything about making your product successful is very dangerous. You’d be sailing blind on your assumptions. And although you may acknowledge you need more facts about e.g. your customers (a known unknown), you should also acknowledge there may be aspects of the problem that you haven’t even got on your radar yet.

That calls for explorative experiments: experiments designed to bring new and unexpected facts to the table.

Such experiments are often more difficult to define precisely, simply because you can't always predict what kind of information you'll find. You're working with rich, ambiguous data, rather than clean numbers. You need to **think like a designer** more than like a scientist.

These experiments are usually Qualitative Experiments, They deal with language, stories, and open questions, and the outcomes are often not (easily) measurable and more difficult to interpret. You still benefit from a well-defined experiment setup, but you're not so much looking to validate or invalidate your assumptions. Rather, you're trying to uncover assumptions and biases.

Qualitative Experiments can help you find out more about the assumptions you're trying to validate. By asking open questions, being curious, and exploring, you can learn more about the world around you. This will lead you to new, hidden assumptions. Often, the findings of an explorative, qualitative experiment help define a subsequent quantitative experiment.

Known Unknowns: Quantitative/Validation Experiments

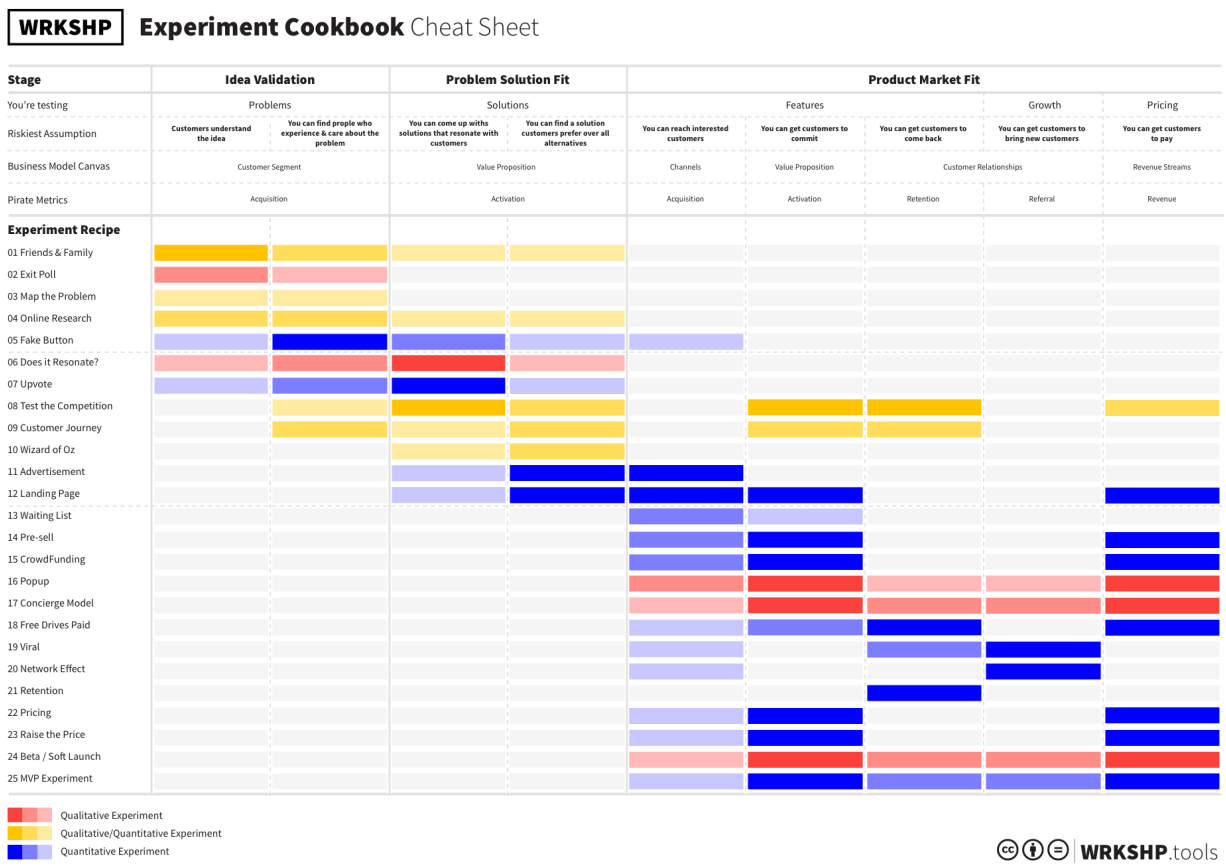
Quantitative experiments give you a clear numerical outcome based on an objective measurement. Think counting the number of people buying a product, clicking a button, or entering a store. Most of the more 'scientifically' based methodology to run experiments, with a clear, falsifiable hypothesis, is geared toward a quantitative approach. To run these experiments, think like a **scientist**.

For quantitative experiments, it's relatively easy to define how and what to measure, and it can be relatively easy to get a clear outcome. The catch is that it can sometimes be difficult to translate your assumption into a simple countable metric. To be able to define a quantitative experiment, you need to know very precisely what you need to know.

You'll need to be very precise and strict about how you design, run, and interpret your experiment to get good quantitative results.

Quantitative experiments deal with the 'known unknowns'. They can't really tell you much about the 'unknown unknowns'.

[4.2] Choose the Right Experiment



(See included PDF)

This infographic shows all of the Experiment Recipes in the cookbook organized by innovation stage, Business Model Canvas building block, and Pirate Metrics stage (see below). It also gives typical Riskiest Assumptions and defines what you are validating (problems, solutions, features, growth, and pricing).

Each recipe is color coded as qualitative, quantitative, or both.

Problem-Market Fit / Idea Validation

During the problem-market fit or idea validation stage, most research will be qualitative. The nature of experiments here is more explorative, because you have less information to go on. You need to cast a wide net and get rich information, and the best way to do that is to talk to people in person.

If you are familiar with the Business Model Canvas, you'll be validating your Customer Segment first in this stage.

Problem-Solution Fit

During this stage, you'll start to also use more quantitative methods. You'll be able to use data and run your experiments with more people, testing your assumptions.

In terms of your Business Model Canvas, you are now looking at the Value Proposition for your Customer Segment.

Product-Market Fit

Although you'll still be running some qualitative experiments (think of User Tests for instance) to get rich information, running quantitative experiments is key in this stage. Data is the most important thing. As the number of (potential) customers grows, you'll have more opportunities to present them with experiments and gather data.

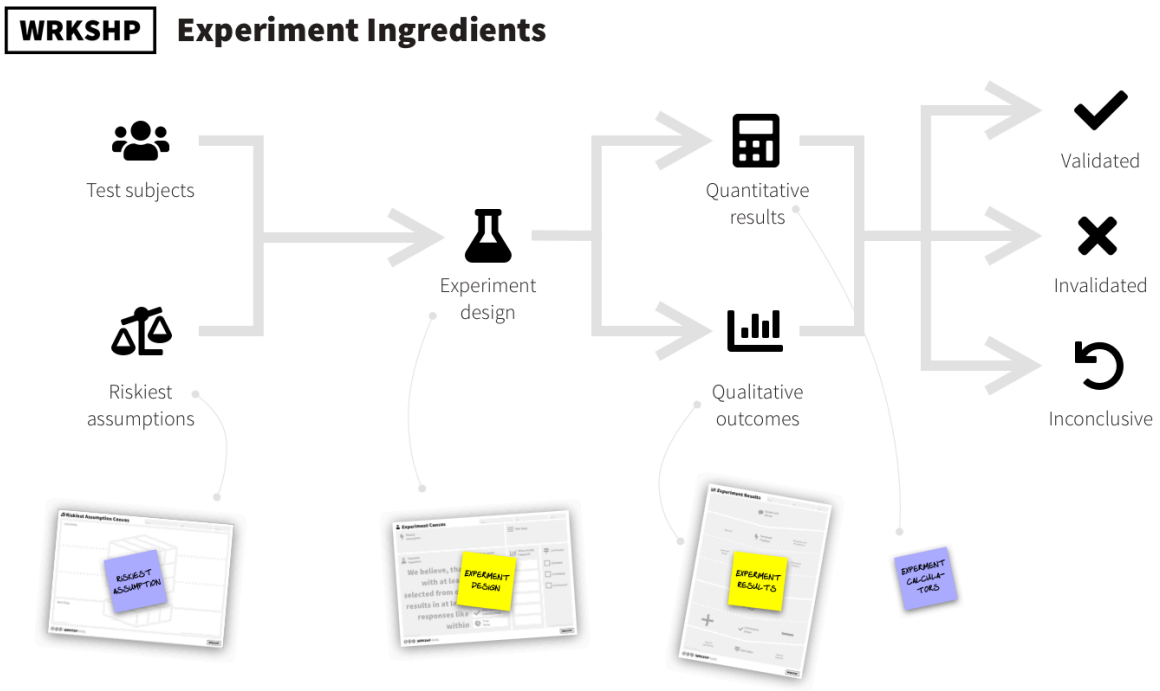
For your Business Model Canvas, you'll validate the Channels, Value Proposition (in more detail), Customer Relationships, and Revenue Streams in this stage.

When you are working towards Product-Market Fit, it can be very useful to use Pirate Metrics as a framework besides the Business Model Canvas. (See: Tools Section of the course for an explanation of Pirate Metrics) In the Pirate Metrics framework, you are validating Acquisition, Activation, Retention, Referral, and Revenue.

[5] Designing Experiments

[5.1] Designing Experiments

In this module, you'll see how to create a well-designed experiment using the 'scientific' method described in module 3. First, let's revisit the experiment ingredients from 3.3.



To **design** an experiment, you'll need to define:

1. If it is a quantitative or a qualitative experiment
2. Your Riskiest Assumption
3. Your Test subjects
4. Your Hypothesis
5. Your Method, prototype, and protocol.

This module describes in detail how to define each of these elements.

[5.2] Quantitative vs Qualitative

The first thing to decide for your experiment, is if it is going to be quantitative or qualitative. Think back to the Rumsfeld quote:

> There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns — the ones we don't know we don't know.
>
> -- Donald Rumsfeld

Quantitative experiments deal with ‘known unknowns’. You can use them to learn more about something you already know exists. Qualitative experiments allow you to explore the ‘unknown unknowns’.

Especially in the Idea Validation/Problem Market Fit stage, you’ll be dealing with a lot of unknown unknowns. Here, qualitative experiments can help you explore. When you uncover a risky assumption, you can follow up with a quantitative experiment to validate it.

Example:
A large Dutch bank wanted to find new ways to use their cyber crime expertise for their (small) business clients. They had no further information about how they wanted to do that, and we had to start from square one. We decided to first run a series of explorative, **qualitative** experiments with small business owners to learn what they thought about cybersecurity, and if they experienced cybercrime as a problem. Once we thought we started seeing a pattern, we converted our assumption into a **quantitative**, online experiment. This allowed us to quickly validate our assumptions.

[5.3] The Riskiest Assumption

Assumptions

Every experiment starts with an assumption to validate. You have an assumption about the world around you that needs to hold true, if your project is going to be a success.

Examples:
Let's say you have the idea of creating a webshop. Some important assumptions:

1. People want to buy what you are selling
2. You can reach these people through an online webshop
3. They are prepared to buy online
4. You can sell your products for a price that allows you to have a profit
5. …

For every startup or new product or service it is possible to come up with a number of assumptions that you will need to prove in order to front-load the risk that is associated with actually building and launching.

In the case of the webshop, it pays off to test the listed assumptions **before** building any complex multi-language international webstores.

In fact, this is what Amazon did. In the 1990s, they had to deal with exactly these assumptions. As Business Insider writes, 'Jeff Bezos had a vision for the company's explosive growth and e-commerce domination. He knew from the very beginning, he wanted Amazon to be "an everything store."' Nevertheless, Amazon started with books, because they had the (proven) assumption those would be easier to buy online than, say, clothing. Only later did Amazon grow into the online shopping walhalla we know today. They started small, first proving they could get customers, before working on better technology and infrastructure.

In fact, in the same Business Insider article, it says, 'Amazon got started out of Bezos' garage and the servers that the company used required so much power that Bezos and his wife couldn't run a hair dryer or a vacuum in the house without blowing a fuse.' How's that for being scrappy!

As an entrepreneur or innovator, you're responsible to tackle the biggest risks that can kill your startup early on. You start out with a vague idea with a lot of misconceptions and unproven assumptions, and you need to learn, you need to build a mental model that is an accurate reflection of reality.

That mental model will always be based on assumptions, there is no way around that. And that doesn't have to be bad: some assumptions may be correct, while others may be wrong, but have little impact on the result.

Correct assumptions and the ones that have little impact are not the assumptions you should waste time on. You're looking for are the assumptions that have a huge impact when they're incorrect. Those are the ones you need to validate.

It's your job to make sure that you find those dangerous assumptions and validate them, the riskiest one first.

Riskiest Assumption

So, how do you find the most risky assumption? Which one should you validate first?

First of all, understand that you are looking for the assumption that is the most risky **at this moment**. Later on, there may be other risks that are even riskier at that moment, but the one you're looking for is the one that dictates a 'go/no go' at this very moment.

One way of looking at it is to think of risky assumptions as 'gating' each other. Some assumptions only become problematic if other assumptions are already proven. But if that earlier assumption is wrong, there won't be a 'later'.

Example:
In the Amazon example above, the quote about the garage using too much electrical power is a clear problem that needs to be solved. Blowing too many fuses at some point meant they would go out of business. But it could only **become** a big problem **after** it turned out they could find so many customers. The assumption 'we can get customers to buy our books online' was more fundamental than the assumption 'we need a better server infrastructure to make sure customers can buy our books online' or even 'we can get thousands of customers to buy our books online'.

The search for the riskiest assumption is not different from first surfacing and then prioritizing your assumptions, and tackling the most fundamental one first.

So, if you would list all of the assumptions that are underpinning your point of view, and rank them from most important to least important, you'd ideally want to start with testing the most important one. In the Lean Startup, this is called 'The Riskiest Assumption'. It's called that, since, if this assumption turns out to be incorrect, you will need to rethink your entire startup.

Luckily, there is a rule of thumb to follow. Each stage in the innovation journey has a number of 'default' associated Riskiest Assumptions. Only when those are validated, you are ready to progress. They are:

- Idea / problem-market fit stage:
    - Customers understand the idea
    - You are able to find customers that experience the problem and care about it (Problem Market Fit)
- Problem-solution fit stage:
    - You are able to find solution options that resonate with customers
    - Customers really prefer your solution over alternatives (Problem Solution Fit)
- Product-market fit stage:
    - You can reach interested customers
    - You can get customers to commit
    - You can get customers to come back
    - You can get customers to bring new customers
    - You can get customers to pay
    - Over 40% of customers do not want to stop using your product (Product Market Fit)

These ‘default’ riskiest assumptions are the ones you should at the very least consider when finding your current Riskiest Assumption. They can often be split up in even smaller assumptions, or be modified to fit with specifics of your project. But if you can’t validate these steps, you’re in trouble.

What about high tech / horizon 3 projects?

In high-tech dependent projects, it often really seems that technology is a factor to test. After all, if you’re developing a new type of algorithm, or you’re finding a better way to generate electricity, or cure a disease, it is quite clear that if your technology fails, you won’t have a product to sell. And proving that the technology works can take a long time, and is really expensive.

Even in these cases, it is smart to use customer development and validation. You need to know how customers will react, and if they care for the problem you’re solving, before you spend millions on a proof of concept. Even if your technology is far out in the future, and people simply don’t ‘get’ it yet, it’s your job to find out what it will mean to your customers to have this new technology, and what they will do with it.

Sure, you can develop a new technology and then simply wait for the use case or the ‘killer app’, but it is a very risky thing to do. What can you do to make it less risky?

Risky Assumptions for explorative experiments

When you're conducting an explorative experiment, it may be more difficult to pinpoint a clear riskiest assumption. Instead, define what you want to learn about, and use that as a guide for the rest of the process.

Exercise

Take the following example startup idea:

> You're going to start a lemonade stand, and dream of having a small shop in the street you currently live in, serving delicious hand-made lemonade to happy customers.

Using the Riskiest Assumption Canvas (see below), try to plot the assumptions behind this idea. Which one is the riskiest?

[5.4] Your Hypothesis

Hypothesis

Your hypothesis converts your riskiest assumption into a statement that can be tested objectively to be ‘true’ or ‘false’.

Given your assumption is correct, it predicts you will see a certain result.

When the experiment is over, you’ll compare your findings with this prediction to determine if the assumption is validated or not.

Creating a clear, measurable hypothesis helps make this interpretation easier and more precise, and doing it **before** running the experiment keeps wiring problems at bay: it becomes much harder to cheat, or fudge the results. (And, who are you cheating but yourself, anyway)

The hypothesis should be quantified. It defines how many people (and from what audience) you need to perform the experiment with, in how much time, and how many of them you predict will respond positively given that the assumption is correct.

To do that, you also need to define when you count people as positive.

All of these aspects have been combined in a template that can be filled in, in the form of the following statement:

- We believe, that (specific testable action)
- with at least (minimum number of respondents)
- selected from our (target audience)
- results in at least (percentage) responses like (what counts as a positive result)
- within (time frame)

To be able to define your hypothesis you’ll need to fill in all of these blanks. The next sections dive deeper into each of those.

What will be your specific testable action?

Specific testable action: This is what you will do as your experiment. It usually ties in with the *prototype* and the *method* of your experiment.

Examples:

- Get people to click on an online ad

- Do face to face interviews
- Conduct a user test
- Pre-sell a product

What counts as a positive result?

What defines a positive result: The type of answer or customer behaviour you count as a positive result (i.e. supporting your hypothesis).

Asking for an opinion is one way of doing this, but it won't give you the strongest signal. As Rob Fitzpatrick writes in 'The Mom Test', people will (sometimes inadvertently) lie to you. They just want to be nice.

The thing is, people have a very hard time predicting what they will do in a hypothetical situation. That is why the question 'would you buy this product?' won't give you any information. It's much better to rely on **actual behavior**. You'll need to come up with a way to link a subject's behavior to a positive result. That can be (preferably) something the subject can do right now (like click on a signup link), or, when that is difficult, a behavior in the past ('how did you deal with this problem the last time you encountered it?'.

In good experiments, the 'positive result' is linked to a **commitment** the customer gives.

This can be small, such as giving an email address, or a larger one such as giving access to network, or actually buying a product or service. The stronger the commitment, the more significant the result, after all, people will be less likely to give you a socially acceptable, 'nice' answer if it costs them something, even if it is as small as running the risk to get spammed.

Examples (from light to strong commitment):

- Like it on facebook
- Pay with a tweet
- Sign up for an email list
- Answer an online survey
- Refer their friends or colleagues
- Buy an existing product
- Support a crowd-funding project
- Sign up for a year subscription
- ...

In one workshop, a team came up with a highly original 'commitment' for their validation subjects: they had to be in a selfie with the team, with a thumbs up. This illustrates that you can be creative with your commitments!

Quantitative experiments

In quantitative experiments, it can be quite straightforward to pick a certain metric or set of measurements, and use that as your positive result.

For instance, one person signing up for your email list would count as one positive result. A visitor to the signup page that does not signup counts as a negative.

Qualitative experiments

In qualitative experiments, it is usually much harder to make clear-cut decisions for a positive or negative. That means, it becomes even more important to define what you're looking for beforehand.

When you're running an interview experiment, for instance, it is possible to specify beforehand (in your protocol) that certain answers or words should be counted as positive.

For instance, when you ask a subject if they have experienced problem X in the last year, when the subject can confidently name three specific situations, that could count as 5 points. One situation as 3 points, a vague 'yeah, i think so' as 1 point, and a negative answer as 0 points. Of course, this is subjective in assigning point values. But it will push you towards something you can at least compare.

How will you select test subjects?

Target audience: The group of people you will select from.

If you are in the problem-market fit situation, you may still be looking for your target audience. In fact, one of the risky assumptions to test is that you can find a target audience of people that experience and care about the problem you want to solve. In that case, you'll need to compare the results from different audiences to see which one is most responsive. You'll probably come up with multiple possibilities.

In the other stages (problem-solution fit and product-market fit), you'll have a much better handle on your target audience.

For your experiments, you'll need to find test subjects from the target group that form a good representation of the diversity you will find in the target group. That means, you should look at the demographics and other data you have on your target group, and double check if you have test subjects representing that data.

Especially in early stages, when it is tempting to use your own personal network as a quick test, be wary of a one-dimensional group of subjects. It's too easy to focus on people that are too

much like yourself in terms of age, gender, education, cultural background… And unless you are making a product that is essentially only catering to yourself, that is quite dangerous.

Control group

If you're really thinking about this from a scientific standpoint, only selecting one group of people is not enough. Let's say you test an assumption and you find out that your target audience, say people aged 55+, like your product, then what is keeping you from running the same test on people under 55? Maybe you were wrong, and your audience is bigger. You can make sure there is a relation between your target audience and your product by also testing a 'control group' at the same time. Doing this randomly (blind or even double blind), and without knowledge of the people conducting the experiment, should also make it much harder to fudge the results. The only problem is that you'll need to do more interviews.

Using pre-existing or supplied audiences

Using the services of firms that supply these specific test subjects, or existing 'customer panels', and other pre-existing groups can speed up the process. But make sure you really understand how your test subjects became part of that group. Are they fans of the brand that will like anything? Are they being paid to take part? These pre-existing relationships may (subtly) color your results.

How many positive results do you predict?

Percentage positive results: The percentage of positive results you predict, if your assumption is true. Above this number, the experiment outcome validates your assumption.

This prediction is the core of the hypothesis. This is what you'll compare your experiment result with.

Benchmarks
To come up with a reasonable prediction, you can use benchmarks. You can look up numbers for other, similar experiments, and compare these benchmarks with your results.

Most quantitative experiments can be split in two groups:

- Single metric experiments
- Comparison experiments

Single Metric
Experiments with a single metric are sometimes called 'Smoke Test' experiments. They measure if a certain signal exists or not.

For example, you could run a ‘fake button’ experiment (recipe 5) and see if indeed people press your button, signaling their interest.

The benefit is that you only need to look for one metric. That means, you will need fewer test subjects than in a comparison experiment. The drawback is that if you find your signal is there, and people e.g. buy your product, you have no information on how well it is doing. Are they more interested in your product than in a competing product? You can’t tell.

Comparison
For a comparison experiment you compare two or more different signals. For example, you could compare two different advertisements, one for a competing product, and one for your own. By seeing which one does better, you get a clearer signal.

The benefit is that you can compare different signals, but this comes at a cost. For each extra option to compare you will increase the number of test subjects needed.

Example: Landing Page experiment
(See: Landing Page Experiment in module 9)
Let’s say you are running a landing page experiment. Your idea is that if you put a new landing page for your product online, and ask people to sign up for a beta test, you’ll be able to tell from the response you have if people are interested in your product.

How can you tell if your landing page is doing well? One way is to see if more people sign up for your product than would be expected for the average product.

To do that, you’ll need to know what a ‘normal’ conversion rate is for the landing page. Your experiment will, in effect, try to measure if your landing page is converting significantly better than the benchmark.

Typical benchmarks for landing page conversion:
- Average conversion rate for a landing page: 2%-5%
- Good conversion rate for a landing page: >10%

Keep in mind that if you are running your experiment with a ‘fake brand’ or a new brand that nobody knows yet, your rates will likely be lower. Also keep in mind, that a lot of optimization is usually required to get your conversion rate towards the 10% range.

Without any prior knowledge, a good benchmark to pick for your landing page conversion experiment is 5% (the top of the average rates). If you do significantly better, your hypothesis is valid. People resonate to this landing page.

How many people do you need?

Minimum number of respondents: You will need some lower bound for your number of respondents. If you fail to meet that lower bound, you won't be able to get any results from your experiment.

This is where statistics come in. Think of it like this:

If you have a large pool of people (the **population**, i.e. your actual target audience), and you're going to interview or test only a small proportion (your **sample**), how certain can you be that the way in which the sample responds can be extrapolated to the entire population?

Example:
Lets say your target audience is 10,000 people, and you talk to 10 of them. All 10 love your idea and immediately sign up. Are you confident that the other 9990 will do the same? What would you bet on it?

Sure, you'll have a strong hunch, probably fueled by learning statistics in school, that you'd like to talk to more people. But how many more, exactly? When can you be confident?

The experiments in this experiment cookbook fall into two broad categories, one where you have very few test subjects (called **low-n** experiments), say 1-100, and a category where you have higher numbers, say a few thousand (let's call this **high-n** here). The first category is often more qualitative, and the second quantitative.

Low-n, qualitative

An example is the 'friends and family' recipe (Recipe 1). You'll ask people from your own personal network a number of open questions. Typically, with a team, you can reach around 50-100 people in this way.

Low-n, quantitative

An example is an 'exit poll' experiment (Recipe 2). You'll find a place frequented by members of your target audience, and ask them to sign up for an email list. Typically, with a team, you can also reach around 50-100 people in this way in a few days of polling.

High-n, qualitative

An example is an online typeform, or a question on Quora. (Online Research, Recipe 4) You may get 100s or 1000s of results, and get qualitative information.

High-n, quantitative

An example is a landing page experiment (Recipe 12) or a waiting list campaign (Recipe 13), where you reach thousands of people through advertising.

Without going into the statistics, there are ways to calculate these numbers.

Example: Landing Page experiment

Applying this to the landing page example, now that you have defined a benchmark conversion rate, we should talk about getting a **significant result**.

How high should your conversion rate be, and how many test subjects should that conversion rate be based on, for you to be **confident** that your landing page performs better than the average, and your assumption is validated?

If you would send your three best friends to the landing page and they all sign up for the email list, you'd have a 100% conversion rate. But it's clear that that would not be a significant result: the three best friends do not reflect the entire population of potential customers. You need more people to go to the page to have a measurement that is more indicative of how potential customers will respond. You have a strong signal, but it's not significant. Even the large difference between your 100% conversion rate and the 5% of the benchmark does not allow you to make decisions, because that 100% conversion rate is a fluke.

If, on the other hand, you'd be able to send a million potential customers to the landing page, you'd get a much more significant result. The variation in the population would most likely be well reflected in the signal you have. The signal might be a lot weaker, and closer to the average benchmark - but it's much more significant. Even a small difference from the benchmark conversion rate would already allow you to make decisions with confidence.

In reality, in early stage experiments, you'll probably end up with a number of visitors that is a lot closer to 3 than to one million.

This can be a problem. How many people do you need to get to the page before you can be confident in the outcome? 100? 1000? 10000?

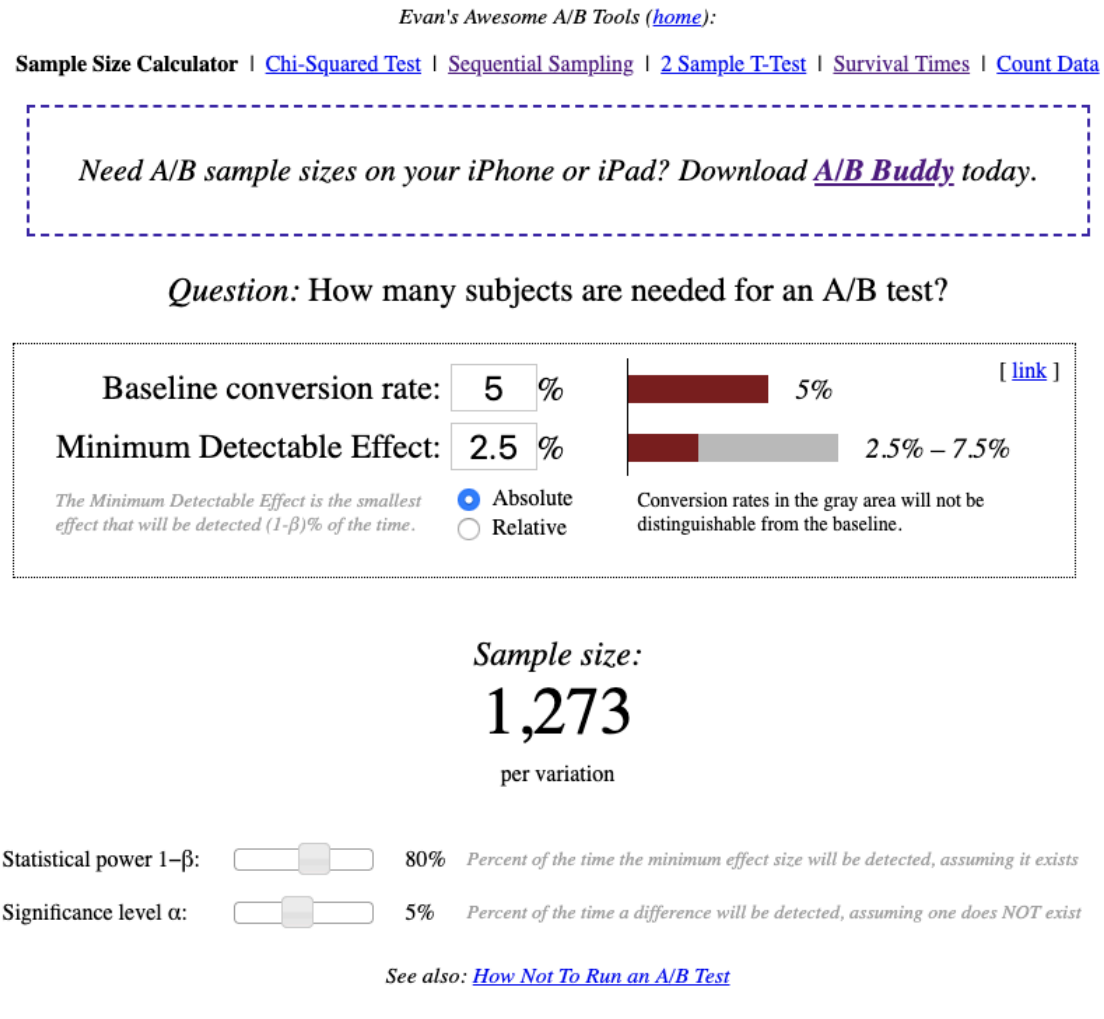
Calculate it

Luckily, it is possible to calculate this number. To do that, it is important to understand that there is a balance between the magnitude of the signal (the difference in conversion rate to the benchmark) and the significance of the result given a number of samples.

If you have a larger signal, you need fewer samples (respondents) to make it significant.

Evan Miller made an excellent calculator that you can use to come up with the numbers you need. [You can find it here](#).

It looks a bit technical, but don't worry, it's easy to use once you get to know some of the statistics terminology.



Screenshot of [Evan Miller's calculator](#)

This calculator answers the following question:

> *Question:* How many subjects are needed for an A/B test?

This may at first seem confusing, as the landing page experiment described above is not an A/B test - it is a test to see if the conversion rate is higher than the benchmark. But when you think about it, this can be easily restated as an A/B test: A is the benchmark, and B is your experiment.

Baseline Conversion Rate

This is where your benchmark goes. I already added in the 5%.

Minimum Detectable Effect

This is the magnitude of the signal you want to be able to pick up.

We said above that we can validate our risky assumption if the landing page performs ‘better than the benchmark’. But how much better? 0.1%? 5%?

The thing is, that there is a balance between the size of this minimum detectable effect and how many test subjects you need to be able to prove that effect exists.

If you want to be more precise, and, say, be able to detect a difference to the benchmark of 1%, you will see you need a much larger sample size. If you want to be less precise, and only count your result as being better than average when you see a difference of 10% above the benchmark, you’ll need a much smaller sample size.

I entered 2.5% here, which means that if your experiment’s conversion rate is below 2.5% (5-2.5) or above 7.5% (5+2.5), you will be able to get a statistically significant result. If your experiment’s conversion rate lies between 2.5% and 7.5%, you won’t be able to tell the difference from the base rate. Any result within these boundaries is too close to the benchmark to be distinguishable from it statistically.

Note: *In early stage experiments, it often makes sense to define larger minimum detectable effects.*

Statistical Power (1-beta)
This is usually set to 80%. It defines the chance your test actually detects the minimum detectable effect, assuming it exists. In this case, there is a 20% chance that it will miss the result, and classify a positive result as a negative result (a ‘false negative’).

It tells you that if you **do measure a positive result**, there is an 80% chance that the outcome was in fact positive, and 20% that you should have had a negative result.

20% may seem like quite a lot, but false negatives really can’t be helped. Trying to increase it to say 95% will really increase your sample size. It is advisable to start with a power level of 80%, and, when you see a signal, to increase the power level in a new experiment.

Significance level alpha
This defines the ‘false positives’. Typically, it is set to 5% (2 standard deviations)

It tells you that if you **do measure a positive result**, there is a 5% chance that it is actually a negative result.

Using the numbers as filled in in the calculator, you would need at least 1273 visitors to your page.

- If you get 1273 visitors, and you measure a conversion rate > 7.5%, you can be 95% confident that you have a positive result and you can validate the hypothesis.
- If, after 1273 visitors, you do not see a conversion rate > 7.5%, you can be 80% confident that you do not have a positive result.

Using this calculator for other experiments

The calculator above can be used for other experiments as well, whether they are qualitative or quantitative. You can use it in the way it was used in the landing page experiment, with a benchmark, or you can use it as an A-B test and compare the results of an experiment on two different target audiences, for instance.

Putting it all together: build your Hypothesis

Now that you have calculated the number of visitors you need and the time needed, you can translate the riskiest assumption in a falsifiable hypothesis.

Example: Landing Page experiment
Using the landing page example from above, putting the hypothesis together looks like this:

Riskiest Assumption:

- We can come up with very different solutions that solve the problem for our customers in a way that resonate with them.
- This experiment tests one of those different solutions, to see if it solves the problem for our customers in a way that resonates with them.

Falsifiable Hypothesis template:

We believe, that a landing page experiment
with at least 1273 visitors
selected from our target audience by running targeted ads,
results in at least 7.5% visitors signing up
within 14 days

[5.5] Method

Now it's time to flesh out the experiment's **method**. What kind of experiment are you going to run? What are the questions you want to ask? What is the prototype you want to show? What exactly is the result you're looking for?

In this stage, you should try to come up with the simplest possible experiment that will give you a strong signal. Really, trim it down. Come up with questions that elicit the clearest response. This can be more an art than a science, so make sure you try out what you came up with and do some **dry runs** (see: Preparing your experiment).

The Method relies on two important parts:

- Protocol
- Prototype

Protocol

The experiment protocol is often overlooked. It has already been touched upon in the hypothesis, in defining your 'positive result'. A protocol specifies this in more detail.

How you handle your data is what ultimately dictates the quality of your results. Setting up quality experiments is a different ball game. While it's probably impossible to remove all bias, it's certainly possible to reduce it. It's not the point to create the perfect experiment. It's more about doing a good enough job and avoiding the big mistakes.

Some basic guidelines:

- Make sure you define your target audience, and **why** they are your audience. If you find test subjects for your experiment, make sure you know for certain that they are part of the audience. **Don't start 'fudging' to get more respondents**.
- Make sure everyone gets the same interview questions, prototypes, and other experiment artefacts, that they are in the same (randomized) order, and that it's clear which ones need to be answered. Are there questions that are optional?
- Make sure you write down location, time, and the method of contact for interviews.
- Define beforehand **how you will score** the output of each question. When does a question count as a positive result? (Remember the points scale introduced above for qualitative experiments)
- Define how to tally up the scores and assign a positive / negative / unclear result.

Following these guidelines isn't that hard, and will get you a long way towards a clear outcome.

TIP: setup a google sheet for you and your team members to use, so you'll capture all the data in one place.

Prototype

Your 'prototype' for your experiment refers to any artefact you need to get a response from test subjects.

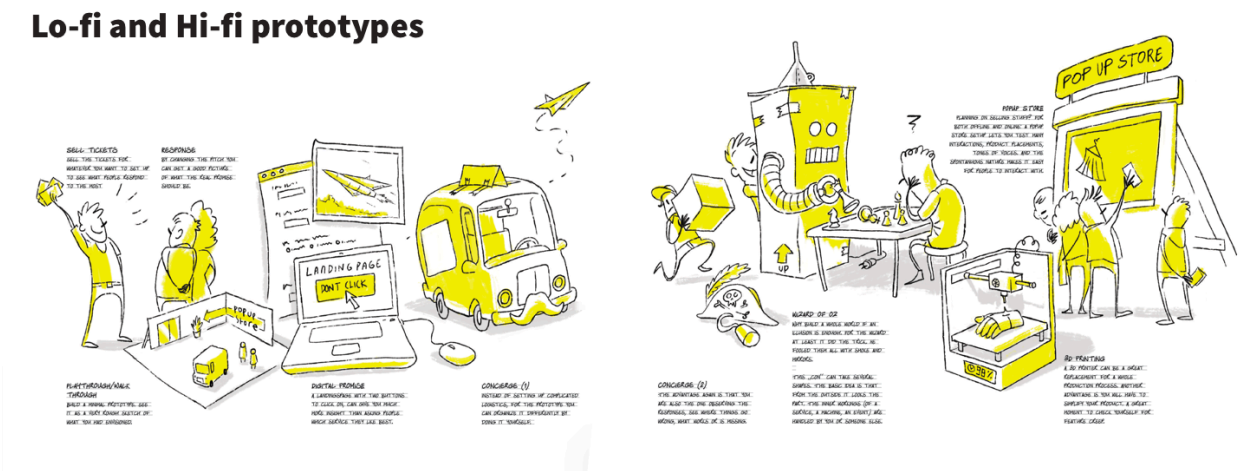
Examples:

- Interview with questions
- A one-page flyer with specifications
- A landing page
- A video
- A rough working version of the product
- Your competitor's flagship product
- …

It usually is something you'll need to create. Try to get away with the most minimalistic version of a prototype you can. The job of the prototype is to get you the experiment results you need, after that it is completely expendable. If it falls apart right after you finish the experiment, then that's fine.

Lo-fi and Hi-fi prototypes

Prototypes can be lo-fi (low fidelity) or hi-fi (high-fidelity). The illustration below (from the book Design A Better Business) shows lo-fi prototypes on the left, and more hi-fi prototypes on the right. Lo-fi prototypes are quick and dirty, easy to make, and very expendable. Hi-fi prototypes take more effort.



For every experiment, you should try to create a prototype that is as lo-fi as you can get away with. It needs to look just convincing enough to get away with it.

[5.6] Your Logbook

When designing your experiment, make sure you write down clearly in your logbook what your hypothesis was, and what the specific benchmarks, numbers of test subjects, etc were that you decided upon. Also, make sure you write down how you made those decisions.

[5.7] Exercise: Reverse-engineer an experiment

To get a bit of practice in setting up a clear hypothesis, I thought it might be interesting to reverse-engineer actual experiments that have been done by startups.

Uber

Uber founder Travis allegedly cold-called ten drivers of black taxis and pitched them his idea. 3 out of 10 wanted to set a meeting with him after this call.

> I went to Google, typed in San Francisco chauffeur or San Francisco limousine, I just filled out an excel sheet and I just started dialing for dollars, right? First ten guys I called, three of them hung up before I got a few words out, a few of them would listen for like 45 seconds and then hung up, and three of them said 'I'm interested, let's meet.'. And if you're cold calling and three out of ten say 'let's meet', you've got something.
>
> —Travis Kalanick, Co-Founder/CEO of Uber

- Write down the riskiest assumption for this experiment.
- Write down the hypothesis for this experiment.

Answer the following questions:

- Is this a qualitative or quantitative experiment?
- In what stage of the innovation journey was Uber at this time?
- What do you think of the riskiest assumption?
- What do you think about the fact that he chose 3/10 as a first metric?
- Do you think talking to more taxi drivers would change the result?
- Was this an expensive experiment to run?

Hely (a Dutch shared mobility startup)

The Hely founding team wanted to know if there was any need in urban areas in the Netherlands to solve a parking problem, and if they were interested in a shared mobility solution. They decided to run an experiment in a newly developed urban area in Amsterdam,

which was built without any personal parking space. They posted 200 flyers in mailboxes with the request to fill out a survey, with the chance of receiving a small reward. The result was that they received 25 filled in surveys, of which 20 were interested in the idea.

- Write down the riskiest assumption for this experiment
- Write down the hypothesis for this experiment

Answer the following questions:

- Is this a qualitative or quantitative experiment?
- In what stage of the innovation journey was Hely at this time?
- What do you think of the riskiest assumption?
- What can you say about the outcome of the experiment? Do you think it is validated?
- Was this an expensive experiment to run?

[5.9] Exercise: Design an Experiment

Using the information in this module, come up with your own experiment. What assumption do you want to test?

If you don’t have a project or startup you’re working on, you can also think of experiments for other companies.

Start from the top, choose a risky assumption, and then go through the module, filling in the blanks.

[5.10] Reading List

Books

- Lean Analytics - Alistair Croll and Benjamin Yoskovitz

Blogs

- Prototype your product Hollywood Style

Tools

- Evan Miller's A/B test Calculator

[6] Preparing Experiments

[6.1] Preparing Experiments

As with many other things, preparation is everything when running experiments.

- Bring together and brief the Experiment Team
- Find and approach Test subjects
- Go over the protocol and make sure any materials (questions, prototypes, etc) are finished
- Do a dry run (preferably with one or two ‘real’ test subjects) to see if they understand your materials and are able to give you the information you need

[6.2] The Experiment Team

To run great experiments it helps to have the right skillset on board. Although not all roles or skills are always needed, making sure you have access to prototypers or designers can do a lot to boost your success. And if you are a small founder team, you'll have to try to fill these roles yourself.

Roles and Skills

Experiment Lead

Always make sure you have a single person in charge of running the experiment and collecting the data. This person does not need to call all the shots, that's not the point: he or she is there to make sure the experiment is run correctly and yields good data.

Data Scientist

If possible, find an actual data scientist to help you make sense of your results, and possibly also when setting up the experiment. If you don't have access to a data scientist, the person with the best head for numbers and science should step up.

Tip: *When you work with a data scientist, be very clear on what you're trying to achieve. Compared to the 'normal' tasks most data scientists are used to, experiments are very messy and vague. Make sure you find a data scientist that can see through this and can be practical, and that is able to make the best of the data they have got.*

Designer (thinking and doing)

Designers are invaluable when running experiments, for two reasons. One is, that you'll need their **design thinking** qualities, especially when you're running early stage qualitative experiments. This is where designers can shine, by really hunting for the 'unknown unknowns'.

Second, you'll need to be able to make quick, nice looking prototypes. This is where the **design doing** side of designers comes in. They are able to create paper prototypes, mockups, flyers, fake brands, and all the other things you'll need.

Prototyper

In some cases you'll need to make something functionally work. In those situations, you may need a prototyper or prototyping developer to help you out

Marketeer / Growth Hacker

Last but not least, having a marketeer or growth hacker on board will help you when you're coming up with new copy, designs, or when you're trying to measure customer behaviour.

Tip: *When you're working with marketeers and sales oriented people, make sure the difference between doing experiments and selling at all cost is clear.*

[6.3] Finding Test Subjects

Finding Test Subjects

It's time to get some respondents together! And, it will definitely take you some time to do this right. One of the most important aspects of running the experiment is where to find your respondents. This very much depends on the kind of idea you have, and the problem you want to solve.

- Qualitative experiments
    - Offline
        - E.g. Interviews, surveys
    - Online
        - E.g. online surveys, chat conversations
- Quantitative experiments
    - Offline
        - E.g. User tests, observations, polls
    - Online
        - E.g. Landing page experiments

(Also refer to the Experiment Cheat Sheet)

The previous module already gave ways to calculate how many test subjects you would need for your experiment. The key is to find people that are from the actual target audience — or what you think your target audience is at the moment — as possible. Friends and family are great to test your interview questions with, but they won't necessarily give you the information you need.

Think about the places (online and offline) where people experiencing the problem gather. What are the channels they use? What are the locations they visit? That's where you'll need to be to interview them.

If you already have access to a channel (e.g. an existing product, a well-visited website, a partner channel) with people from your target audience, all the better. In fact, for many startup ideas, not having access to people to interview and run experiments with is a real drawback.

Sources for offline test subjects and subjects you'll interact with personally

Source	**Pro**	**Con**

Your personal network and linked in	Easy to reach	Biased, probably less diversity, maybe no good match with target audience, smaller group
Friends of friends	Easy to reach, slightly larger group,	Gated by friends, still maybe no good match with target audience
Existing customers and relations of the company	Easy to reach if you already have a business	(Perceived) risk to current relationship, biased
Focus groups and customer panels	Relatively easy to reach	May be biased, not a typical user (trained as focus group member), paid for their opinion, small group
People you approach at meetups or conventions	Pre-selected by interest, which can make it easier to find people from the target audience. They are there to meet people.	A lot of competition for attention, they may be in 'polite mode'.
People that respond to an ad or announcement	You can pre-select them, making it easier to target a specific group	Ads can cost money, you won't always know what you get.
People you approach at suitable offline locations	You can find places frequented by the target group, selecting	People are often in a hurry, suspicious of sales pitches, not in a mood to talk, often a mix of people outside of the target audience

There are many more possible sources you can think of. Before you pick one, think about the pros and cons. And if you do pick one: stake out the venue. Do a dry run, before you commit your resources to do a full scale experiment. Experiments where you interact with people face-to-face are often very rewarding, but they are also expensive, since you will have to perform them yourself.

Sources for test subjects won't interact with directly

Source	**Pro**	**Con**
Visitors on an existing website or online channel	If you have a good channel, you'll have lots of traffic to work with, it can be fast	You need to have a channel, you need to get permission to run your story, ad, or link on

		that channel
Email lists	If you have (or buy) a good email list, you can reach a lot of people quickly	Email lists can be expensive, using a bad list (not the right audience) will give you bad results, email click rates can be low
Content marketing	Posting specific content can attract the right audience	It takes a lot of time (both making content and getting results)
Guest posts / PR	You can find the channels that have the right audience, it can be quick	Unless you're close friends or partners, this may be expensive.
Flyering / handouts	You don't need to know email addresses or other personal data, it's geographic	Old school, low response rate
Google Search ads	Using search keywords can give you quick ways to validate and compare, quick to roll out, cost can easily be capped, and you only pay for clicks (the positive results you measure), great analytics data	It costs money
Most other advertising	Can be seen offline, targets specific areas	You pay for exposure (so even if nobody responds you still pay), old-school, sometimes low response rates, no analytics data

Tip: try to define an experiment around a channel where your target audience is already present. Offline that could be a physical location, online it could be a specific website, email list, or forum. Do you have access to an existing channel? Perhaps through partners, existing products? Then that will be a good place to think about for your experiment.

Checking Subject Suitability

Make sure the test subjects are from your defined target audience, and double check if you have enough diversity. If you defined an age range from 20-50, and you interview only 20 year olds, that will skew the results. Also, try to find more people than you need, since there are bound to be a few no-shows.

Tip: *When you design your experiment, try to add in some questions to check suitability. Ask for demographic or other information if that helps confirm the subject is part of your designated target group. For example, if you're targeting people that love to listen to music, ask them to estimate how much time they have spent listening to music the last week, rather than if they 'love music'.*

What does my experiment cost?

Now that you know how many test subjects you need, is it worthwhile for you to run this experiment? This really depends on where your traffic comes from.

Advertising traffic

- If you already have a source of traffic available and you can redirect some of it, e.g. by placing a link, sending an email, or writing a blog, it is definitely interesting and cheap.
- If you don't have a source, you may need to advertise to reach traffic. You'll incur costs depending on the type of business you're in and the keywords you'll need to use. You'll need to calculate the cost beforehand.

Example:
Using the calculation example from the previous module, for a landing page experiment, you have the following numbers:

Traffic Needed: 1273 (minimum)
Cost per Click (CPC): $2.69 (benchmark from Google ads)
Cost: 1273 * $2.69 = $3425
Clickthrough (CTR): 3.17% (benchmark from Google ads)
Reach: 1273 / 3.17% = 4K

The traffic needed comes from the A/B test calculator. The Cost per Click (CPC) is what Google Search ads charges you each time someone clicks on your ad (ending up on the landing page). Cost can be calculated from the required number of visitors times the CPC.

Note: *The clickthrough and reach are only shown to give an indication of how many visitors will need to see your ad to get to the required number of visitors. It can help to calculate that reach to see if it far exceeds your target audience size.*

As you can see, it will cost you around $3425 to run this experiment this with advertising.

Note: *the CPC rates vary heavily depending on the industry you're in. Look at the tables below to get more insight*

For early stage landing page experiments, therefore, it is often preferable to use traffic coming from existing channels. Build that email list, comment on Reddit or Quora, strike partnerships, put your idea on Product Hunt. Be creative. Be scrappy. Find the places online where your potential customers are, and find a smart, creative way to get your landing page on their radar so they come to your page for free.

For later stage landing page experiments, especially when you already have a product to sell and you're measuring sales conversions rather than signups, it can be worthwhile to use advertising – but do the math before you start.

Average Clickthrough rates (CTR) for Google Search Ads per Industry (2019)

Industry	Average CTR	Average CPC
Advocacy	4.41%	$1.43
Auto	4.00%	$2.46
B2B	2.41%	$3.33
Consumer Services	2.41%	$6.40
Dating & Personals	6.05%	$2.78
E-Commerce	2.69%	$1.16
Education	3.78%	$2.40
Employment Services	2.42%	$2.04
Finance & Insurance	2.91%	$3.44
Health & Medical	3.27%	$2.62
Home Goods	2.44%	$2.94
Industrial Services	2.61%	$2.56
Legal	2.93%	$6.75
Real Estate	3.71%	$2.37
Technology	2.09%	$3.80
Travel & Hospitality	4.68%	$1.53

Source: https://www.wordstream.com/blog/ws/2016/02/29/google-adwords-industry-benchmarks

Email lists

Another often heard approach to quickly access a large number of test subjects is by using email lists. While this can be very successful, there are a few things to look out for.

Buying vs. Renting
Buying a mailing list means that you own the list, can use it as many times as you want, but you are responsible for keeping it up to date.

Increasingly common, however, is mailing list rental. With rental, you pay for one-time access to a list or unlimited access within a specified time period. It's cheaper, on average, to rent a mailing list, and some companies only allow you to rent mailing lists. You're not responsible for keeping the list up to date, and you won't have to deal with GDPR and so forth.

Compiled lists vs. Response lists: Compiled mailing lists, also known as "cold lists," are pulled from databases with consumer and business information obtained from numerous sources. The contacts on a compiled list did not necessarily respond to previous direct mail requests, and response rates for these lists are generally lower. There may be a lot of 'garbage' on such lists.

Response lists, on the other hand, consist of people who have previously responded to direct mail requests or given permission for companies to contact them. As a result, you can expect a greater response rate from mail sent to these customers.

Business list vs. Consumer list: For a more targeted (and successful) experiment campaign, you'll probably want to narrow down your list to include either businesses or consumers (though most brokers can provide both).

Business to business mailing lists are put together from data sources such as business directories and government agencies. You can further narrow down a business mailing list by using selects that include number of employees and annual revenue.

Compiled from phone directories, birth certificates, and other sources, consumer mailing lists can be narrowed down by using selects like credit rating and household income.

Something to keep in mind when buying mailing lists:

List quality: The success of your campaign is highly dependent upon having a good list. With mailing lists for marketing, like most things, you get what you pay for. **Extremely inexpensive lists are likely to be out of date or poor quality**. Lists filtered by multiple, specific criteria, while pricier, should provide a better experiment result.

The cost to buy a mailing list is typically based upon the cost per thousand impressions (CPM), although you can also evaluate it on a cost per name basis. This means

- For a consumer mailing list, expect a CPM cost of around $50 to $200 ($.05 to $.20 per name).
- For business mailing lists, the CPM is around $150 to $350 ($.15 to $.35 per name).

(This section was adapted from this CostOwl article)

Note: *The price can differ depending on the filtering you require for the list. Count on a list of CEO email addresses (if such a thing exists) to be much more expensive than a random list of people.*

Example:
If we use an email list to get people to the landing page in our landing page experiment example, we get the following numbers.

Samples Needed: 1273 (minimum)
Clickthrough (CTR): 2.5% (benchmark from Campaign Monitor)
Addresses Needed: 1273 / 2.5% = 4920
Cost per address: $0.20 (benchmark from CostOwl)
Cost: 4920 * $0.2 = $984

Other sources and additional costs

When you're using different (online or offline) sources, it still makes sense to calculate what the experiment will cost you before you run it. Think of things like:

- Time needed in interviews
- Renting rooms
- Food & drink for people when they conduct interviews
- External equipment you may need to rent (cameras, microphones, etc)

Also think about other costs:

- The cost of making a prototype
- Design cost
- Printing cost

Use these calculations as a factor in you decision to run a specific experiment. Understanding the cost involved will also make people more alert to running the experiment the right way.

[5.3] Fine tune your protocol and materials

Once you have come up with a good source of test subjects, you'll likely need to revisit your protocols and materials.

Do you have everything you need to approach the test subjects? Depending on your experiment and your location, maybe you decided to use advertising, a landing page, a sign, flyers…

For offline experiments, is everyone instructed? Do they have a time table, a map of where to show up, etc?

Are the questions you want to ask, or the activities you want to measure, clear? If you are running an online experiment, are you sure your email addresses are being gathered? Is your analytics setup linked to the correct accounts? It would be super sad to see an experiment go to waste because you forgot to turn on your analytics, and have measured nothing...

[5.4] Do a dry-run

Maybe the most important tip: do a dry run. It may seem tedious and a waste of time, but it is absolutely vital.

For interview experiments, do some practice runs. Go over all the questions with a real potential test subject. Do they get it? Is everything clear? It's completely human to be biased in the questions you ask, and doing this in pairs will help you spot the most obvious ways in which you are subconsciously influencing the result.

Show your online advertisement or landing page experiment to potential test subjects. What is their reaction?

When you've finished preparing, it's time to start!

[7] Running Experiments

[7.1] Running Experiments

Running online experiments is, usually, well, running the experiments. In case of an online experiment, there won't be too much to do, except monitor your analytics and checking if everything is recording properly.

When you're running an experiment where you have to interact with test subjects, it's a different story. You and (hopefully) your team will be hard at work approaching test subjects, getting them settled in, and asking them questions. In the mean time, you'll need to record what is going on.

When you're running these kinds of experiments, be careful with your team and your energy levels. Don't push too hard: tired people make bad interviewers. Sure, it's cool to get more interviews in, but in the end you're looking for quality data. Do a check in regularly to see how you're doing, and don't schedule too many.

In both cases, as long as you see data come in, and that data looks solid, you've got nothing to worry about. If you suspect your data is not solid, you should prepare to step in and fix the problem.

Several situations might occur.

1. You are not getting test subjects or visitors at all. The source you targeted is not delivering.
2. You are getting test subjects and results, but the results are all positive or all negative. This can mean two things:
    a. Test subjects are invalidating your assumption
    b. Test subjects don't understand or for some reason ignore your experiment.
3. You are getting test subjects and results, but the results are all over the place and seem random. Again, this can mean two things:
    a. Your experiment is unclear, and you're not asking the right questions or presenting the right prototype. People are responding based on their own interpretation.
    b. There actually is no clear answer, and you'll have to go back to the drawing board.

If you get a 100% positive or 100% negative response, that is usually also a reason to take a hard look at your experiment. You may be asking an obvious question…

Some of these problems will be ironed out when you're doing your dry run (see previous module).

[7.2] Recording Results

Recording Data

Ok, so now you're running your experiment and you're getting data. It pays off to think beforehand about where you will store the data and (again) to do a dry run to see if you are getting all the data you need. It's just not cool to figure out that you've been asking the wrong question after you have just spent 3 days interviewing.

Tip: *Storing your data with your team in a common Google Sheet that already has all the questions (and ideally, how you'll score them) is a great way to capture both qualitative and quantitative data.*

Logbook

Record data in a logbook or your google sheet. Try to capture the time, who was present, and what went on in the interview or during the observation. Write down your observations of the test subject's mood, the way they answered, and other peculiarities in your logbook, but separate them from the recorded answers.

When you're running online experiments, and analytics are being kept automatically, capture the moment the experiment starts, stops, and try to get some intermediate results, so you tell if your experiment is running correctly. Once the experiment stops, download the data you have generated.

Make sure to also write down the times and dates of important activities during the experiment, e.g. when you sent out mailings or when you adjusted your experiment for some reason, and why you did that.

You'll be surprised how fast 'data rot' kicks in when your experiment is finished, and how difficult it is to reconstruct the exact timeline even a few weeks later. Give yourself a break and use a logbook for that.

GDPR and Privacy

Check if your data collection complies with GDPR or privacy rules. If possible, anonymize your data, assigning test subjects a number. This also helps with preventing bias. When you're interviewing people, make sure you ask their consent, especially when you use photography or video.

[7.3] What to record?

When you're recording data, make sure you try to record the answers as complete as possible. Writing down your own thoughts or interpretations won't help you. Try to capture what the test subjects are saying or doing, and keep your own thoughts separate. It is too easy to start interpreting when you should be observing and measuring. That way, you'll definitely have a lot of confirmation bias.

The 'Experiment Result Canvas' in the next module has some great categories of things to look for while interviewing or observing.

[7.4] How to record?

When you are recording data, be smart about it. In interviews, have one interviewer ask questions, while another records the answers and observes. Alternatively, use technology and record the interview (your phone probably has a great memo recording tool). Skype or video call conversations can easily be screen recorded, for instance using Quicktime or Camtasia. When the interview is over, you can go through the results at your own speed, or, still better, use an automatic transcription service and get instant written text (for instance using Trint, where you can get unlimited transcriptions for around $50)

The point is: try to use technology to help you, so you can focus on asking questions and getting answers. If you have to do both, you're going to slow down, skip recording important information, or, worse, you get carried away in the interview and end up having to fill in the answers from memory.

[8] Interpreting Experiment Results

[8.1] Interpreting Results

Your experiment has finished, you have gathered your results, and you're ready to find out if your assumption was validated.

Now comes the hard part: interpreting your data.

Even if you have a very straightforward, quantitative experiment, in many cases you'll find out that drawing a solid conclusion can still be a challenge.

Why?

Ambiguity and uncertainty.

In many cases, especially with qualitative experiments, you'll see ambiguous results and conflicting evidence, that you can interpret in multiple ways. It won't be a clear cut case. And even when you get a seemingly clear result, you may still feel uncertain about taking it as a guide for your next step.

Basically, it comes down to trust: can you have confidence in your result? It's a tough question to answer, when it has potentially life-changing consequences for your startup and the founding team.

If your customers tell you they love your product, measured by a high conversion rate, can you safely lean back and think you've solved the entire puzzle? Or will you get a nagging feeling you may have underpriced your product?

There will always be ambiguity. Even if you can prove your product works by hundreds of sales, there are still lots of things you don't know. The feeling you can improve something will stay. And that's ok. It's part of business.

Running experiments is not about getting rid of this ambiguity and uncertainty. It's about proving just enough to warrant your next (small) bet.

That's why you need to work up your experiments from Idea Generation, through Problem Solution fit, towards Product Solution fit: every step reduces your uncertainty enough to warrant the next step. You're not betting the farm all at once.

This section deals with interpreting your data and making a decision, and dives into what to do when your data doesn't give you a clear answer.

[8.2] Clean your data!

The first step towards a conclusion is to look at the data you have gathered. Put it all together. What have you got? Double check if you have the latest versions and got all contributions.

When you've collated everything, it is time to look at the answers in detail. You'll need to **clean** your data.

When you're dealing with interview results, look for skipped questions, answers that were accidentally misplaced, or that you can't make sense of. You'll need to either correct mistakes if it is obvious how, or you'll need to remove bad results. You might talk to the interviewer in case something is unclear, maybe they can give pointers as to what happened, or why a question was skipped.

Did all interviewers record the data in the same way? Be wary of things like 1-10 scales, where people may have used 10 for 'high' and others have used 1. You get the picture. You won't get away from reading through all the results.

When you're dealing with quantitative data, you should still take a hard look at your data and correct any problems, before going further.

For instance, if you have been running a landing page or other online experiment and you have been counting page views, you may have been counting your own (test) page views as well. You should have taken care of this when preparing your experiment, but it happens that some of these false readings were missed. You'll need to correct for this, especially if you have a low-n experiment or if you have been trigger happy testing your own site during the time the experiment was running. If you're looking for 100 hits, and 30 are your own, that's a big deal. If you're looking for 50 signups, and you have added 10 test signups, your results are 20% off. And even if you add 10 test signups on 1000 email addresses, that is still a 1% difference.

Besides looking for your tests contaminating the result, you should be looking for other suspicious behaviour. Did you get lots of fake-looking email signups? Maybe a bot found your signup form which was not captcha protected. Did you get strange traffic spikes where your conversion rate suddenly dropped to zero? Same story. Try to assess if your data is solid, and try to exclude fishy results.

Cleaning all this data may mean you have not enough solid data left. That can be a problem, and it may mean you need to continue your experiment for longer. It helps to always take into account some data cleaning loss when setting the minimum numbers of test subjects in your experiments.

[8.3] Interpreting Qualitative Data

Interpreting qualitative data can be difficult. There is a lot of ambiguity. First of all, you need a way to collect and share the information you received with your team. And, even more important, you need a way to draw conclusions from the results.

One method that is used a lot to make sense of experiment results in a workshop setting is to use dot-voting (Dot voting explained on Wikipedia). To see how this works — or actually, totally doesn't work — here is an example.

(Note: Dot-voting itself is arguably a completely flawed method, but that's not what this article is about)

Example

A startup wants to develop a new app that will help self-employed people make sense of their finances. They just started out, and are in the early stages of their journey.

Although initial responses (from their own crowd of friends and acquaintances) are positive, they have barely begun formulating the problem they want to solve for their customers.

They are at a crucial stage for their startup: they are hard at work trying to move from a mindset where they just want to build the thing that is in their head, to something real customers actually want to pay for.

Following the lean startup, the best way to make this mental leap is to confront the ideas living in your head with real people out in the real world. Following this advice, the team goes out on an exercise where each team member interviews 10 self-employed entrepreneurs. They ask them how they deal with their finances, and how they keep tabs on things like their cashflow and invoicing. When they come back, they together have gathered around 50 interview results.

To save time, they decide to print the results, stick everything on a wall, and then go over it as a team. They plan to look at all the results and then use dot voting to mark what they find interesting.

On the surface, this looks like a valid approach. Reviews like this can work great in a workshop setting, where the information is on the wall, clear for everyone to see. The dot voting process is fast and gives you a clear result.

This approach is broken!

However, when I used this and similar methods in the past to make sense of qualitative results, I noticed that it **really doesn't work**.

People tend to focus on the interviews they have conducted themselves. They notice the things that they already agree with, or have noticed before. This leads to a **strong confirmation bias**. With this method, you'll most likely reinforce any bias that was there to begin with.

Clearly, what's needed is another approach. But there are some constraints.

Initial interviews are necessarily exploratory in nature. At this point, the team can't really know in detail what they are looking for. There won't always be a clear script to follow, and the answers are varied.

You simply can't follow the approach that would be used for a large scale survey and use statistics. The low number of responses and the unstructured nature of the results make that impossible. This makes it very difficult to make a clear 'validated or invalidated' decision.

From a statistical point of view your low-n results may be a complete waste of time. But at the same time, they are a treasure trove of information about your customers.

How to fix it

So, what can be done? How can we explore qualitative interview results in a meaningful way? A way that is as objective as possible? How can we extract as much useful information from experiments as possible?

This section describes a number of steps you can take to ensure you can take confident decisions based on your experiment data.

Exploration vs validation experiments

As was discussed earlier, qualitative experiments are often used for exploration purposes, without a clear assumption to validate. At other times, they are run as a normal experiment and you want to score results so that you can validate or invalidate your assumption. Before you start interpreting the data, remind yourself of the purpose of the experiment.

Exploration experiments

Many early stage experiments for Idea Validation or Problem Market fit are at least partially explorative in nature. You want to get a big picture of how different members of the target

audience think and behave about the problem you're interested in. You want to find out if there are clear patterns in their answers.

Finding patterns
One of the best ways to find patterns in data is by visualizing the data. Using e.g. Google Sheets, it is very easy to turn answers into bar charts or pie charts, giving you an idea of different groups in your audience. Of course, there is also specialist software that can help you do that, but it often requires specialist knowledge.

Quick & Dirty Graphing using Google Sheets
With Google Sheets, it's very straightforward to turn your cleaned data into nice charts that help you see patterns in a glance.

Start by putting all your questions in rows, and your test subjects in columns. The answers for each question now line up.

	Subject A	**Subject B**	**Subject C**
Question 1	Answer A.1	Answer B.1	Answer C.1
Question 2	Answer A.2	Answer B.2	Answer C.2

Tip: *Use a sheet with this setup from the start.*

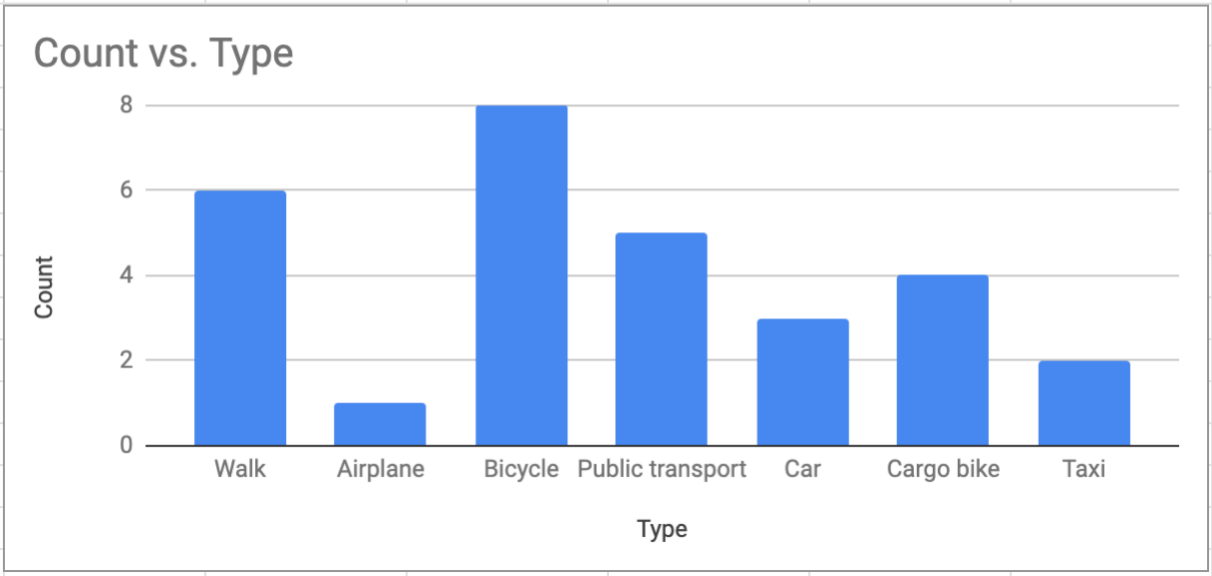
Doing this will help you to easily compare different answers for different subjects.

Next, define scoring per question (or use what you have defined in your experiment protocol). Turn each answer into a score.

For instance, let's say you did an interview to learn more about the different types of transport people in your target audience use. A question was: 'Which types of transport did you use in the last month?' Answers were for instance 'Bicycle, car, walking', or 'public transport'.

First, go over all the answers and make a list (in a new sheet) of all the different types of transport that were mentioned. Now, you can go over the answers and score them for the types of transport they mentioned, as well as the number of types per question. Tally up the results per type of transport. In the new sheet, you can use 'insert chart' to plot these results.

Type	Count	Notes
Walk	6	
Airplane	1	
Bicycle	8	
Public transport	5	(Bus, train, and other forms combined)
Car	3	
Cargo bike	4	
Taxi	2	(Taxi, uber, and other services)
Total	29	



Example: *bar chart for explorative data. Looking at the numbers makes it less obvious, but the graph clearly shows that apparently, in this survey, more people use cargo bikes than cars. If that is different than you expected, it may be interesting to follow up on.*

Ways to score data

Count specific words: define a list of terms or words that are relevant to what you want to know, and tally if or how often they are used by a test subject.
Example: We did an interview about how people organize their personal finances. When people would volunteer things like ‘savings account’, ‘investing’, and other relevant terms, we would tally those. It turned out some people seemed to use more ‘risk averse’ words like saving and security, and others were more ‘aggressive’ and used words like ‘invest’ and ‘growth’.

Look for specific numbers: for some questions, it makes sense to note down a number.

Example: For the same personal finances interview, we asked people how often they would check on their finances, and what was the last time they did it. They would either give a specific number, say ‘every month’, or ‘every couple of weeks’, or, in some cases, they would only be able to say ‘oh I looked at it last february’. We used this information to score in buckets of ‘every week’, ‘every month’, ‘a few times per year’, and ‘rarely’. With these, we could then again make a visualization.

Positive and negative answers: For many questions, you’d like to convert a longer answer into a yes/no outcome.
Example: Again in the same interview, we asked test subjects if they had ever invested in stocks. Some would give a long answer with why they never did that, or were thinking about it, and others gave a clear yes/no. We translated all of them to yes, no, or unknown, and counted the occurrences.

Tip*: go over the same question for all the results, instead of scoring one result at a time. This promotes uniform scoring.*

There are many more ways to score results in this way. Just remember, that with any scheme you pick, you’ll be adding a layer of interpretation on top of the raw data. When you’re looking at the resulting visuals, try to find outliers, look for combinations of patterns, and try to see if you can find things that conflict with your assumptions.

Ask yourself questions such as “Are the people that use more ‘risk averse’ language talking about their personal finance also the ones that check their finances more often?” (What do you think the answer to that question was?)

Experiment Results Canvas

To make the process of going through results as a team a bit easier, I’ve added the ‘Experiment Results Canvas’ below.

Validation Experiments

Scoring for qualitative validation experiments can be done in the same way, by assigning values to answers. The difference is, that you should look at your hypothesis and assumption: they should give a prediction of the outcomes.

For example, if your assumption is that people in your target audience experience the problem of finding good parking spots in the city center, and in your hypothesis you say this assumption is validated if at least 30% of test subjects agree, then you need to convert answers to the questions in your interview to match that scale.

The question ‘describe the last time you had to park your car in the center’ might give rise to an answer that goes into detail about the horrors of a 45 minute search for a parking spot. This we can definitely count as positive.

Once you have tallied up the scores, you can directly check against your hypothesis.

[8.4] Interpreting Quantitative Data

Interpreting quantitative data is often a little easier, since the scoring is already done for you, and you can rely on your strict hypothesis from the experiment design.

Still, sometimes different data sources need to be combined to a score. Try to figure out how to match what you have measured (e.g. clicks on a landing page) to your hypothesis.

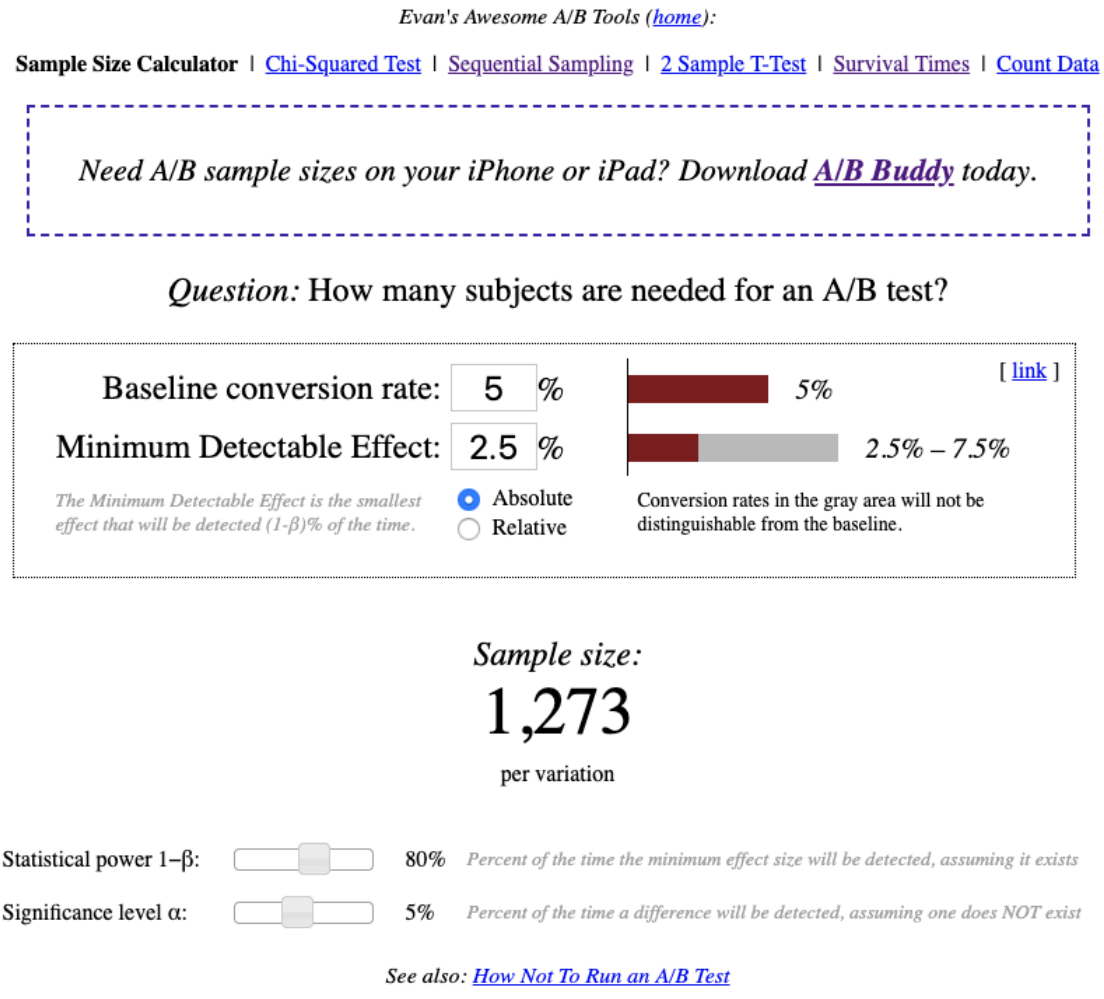
[8.5] Making decisions

This is the step you've been doing all the work for. This is where you find out if your assumption is validated or not.

As mentioned in the introduction of this module, the goal is to be able to decide with confidence.

How to make a confident decision?

The heavy lifting for your decision has already been done in the experiment design. Referring back to that section (4.3), you defined a clear hypothesis with a minimum number of respondents, and a minimum effect to measure.



Remember Evan Miller's calculator?

In the example in section 4.3, the hypothesis that was created looked like this:

Falsifiable Hypothesis template:

We believe, that a landing page experiment
with at least 1273 visitors
selected from our target audience by running targeted ads,
results in at least 7.5% visitors signing up
within 14 days

The good thing is, that by using a calculator and setting a strict hypothesis, all you need to do now is see if you have met the criteria, using your cleaned data and your scoring.

In the case of the example, there can be multiple scenarios:

- You have achieved a conversion rate that exceeds 7.5%. You can validate your assumption (with 95% confidence).
- You have achieved a conversion rate that is lower than 7.5%. You can invalidate your assumption (with 80% confidence).

What to do when you can't make a confident result?

Let's say you have some doubts about your result. You've run the landing page experiment from the example and you've come up with a conversion rate of 7.4%. It's lower than 7.5%, so technically you should invalidate your assumption, but it's so close! Even if the math behind the calculator tells you the result is significant and that you should be confident, this number doesn't make you *feel* confident. What to do?

If the results are this close, the easiest thing to do is to extend the experiment a little longer. The more samples you have, the more the measured conversion rate should start to resemble the true conversion rate. Another thing you can do is simply re-run the experiment, if possible with more test subjects. It will cost you some extra time, but if it gives the same result, you know where you stand.

If you get a vague or ambiguous result from a qualitative experiment, try to simplify the experiment and re-run it. Rather than having 10 open questions, try to turn the question you're most interested in into something you can test quantitatively.

Now, let's take a situation where you've run your experiment, and predict that 30% of all test subjects should be positive to validate your assumption. To your surprise, you get a result of 0%. Not a single test subject likes the idea. Of course, technically you should reject the result, but it seems fishy. You can't be that far off the mark, can you? What to do?

The first thing to do is look at your experiment setup and your data, and see if you can think of a reason the experiment gave you this outcome. If the results are really off, try to get in contact

with some of your test subjects and ask them how they experienced the experiment. Do they understand what you want them to do?

[8.6] Pivot or Persevere

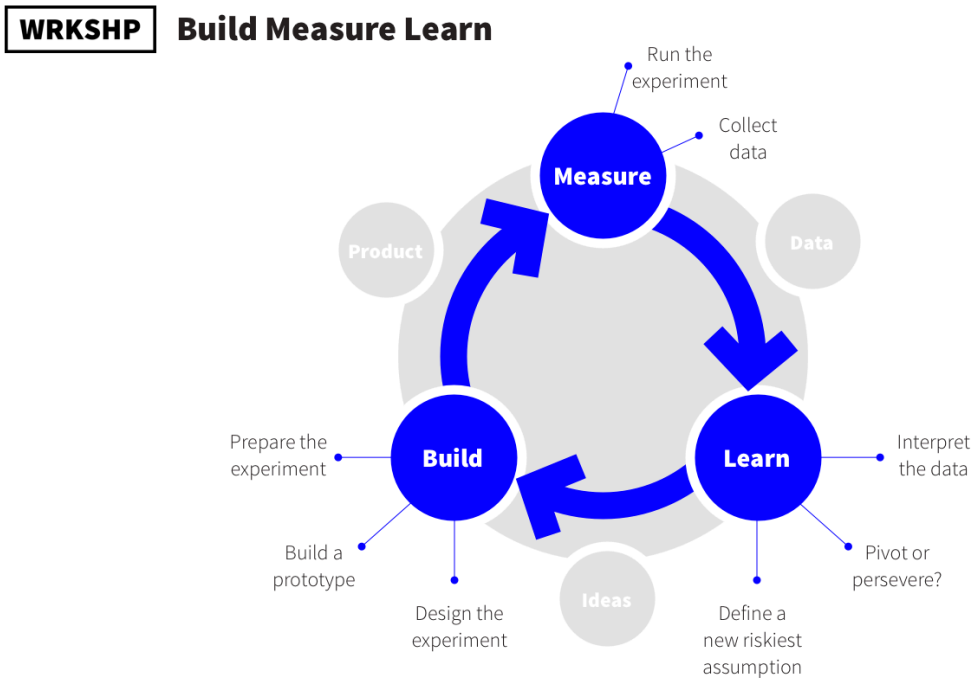
Pivot or persevere

After coming up with a conclusion, it's time for your next steps. If your riskiest assumption was invalidated, you need to do some hard thinking and really go back to the drawing board. You'll need to Pivot, and go back to your design criteria, and come up with a different solution.

If your assumption was validated, it's relatively smooth sailing. Simply move your current riskiest assumption to the 'validated' pile, and try to find the next riskiest assumption with your team. And based on that, define your next experiment! Repeat until revenue.

> Breakthrough insights are often hidden within failed experiments.
>
> — Ash Maurya

[8.7] Closing the Loop: Running Multiple Experiments



Section 3.4 already described the build-measure-learn loop and how to string together multiple experiments. Iterating your experiments, and using a string of experiments to move you from idea validation/problem-market fit to problem-solution fit, and from there to product-market fit, is at the core of the innovation journey.

Doing this effectively means that you need to keep track not just of your current experiment, but you also need to remember what you have already validated. Your logbook is invaluable for this, as is the Riskiest Assumption canvas.

Keep track of experiment outcomes, why you did these experiments, and what you think your next riskiest assumption is.

Tip: *one thing that can help you tremendously while going through the early stages as a startup, is to start building an audience of potential customers. Many startups don't have a good source of traffic and test subjects to validate with, and end up having to pay for their experiments. If you have your own audience, you always have access to the opinions of your potential customers. Start building an email audience immediately, and keep them informed of progress.*

[9] Experiments for Idea Validation

[9.1] Idea Validation

Experiments for Idea Validation

In the first stages of developing a new startup or product, you'll need to make sure that there is an actual problem that your idea solves.

Popular terminology calls this stage 'idea validation', but you've got to be a little careful with that term. 'Idea validation' opens the door to confirmation bias, as it implies you're looking to see if your idea is 'correct'. In my experience, early ideas almost never are. Ideas start out rough, lopsided, undeveloped, and based on a very limited experience (your own).

Would you rather be correct? Or successful?

It isn't really interesting to see if your original idea is 'validated', either. It is much more interesting to see if your idea connects with a **problem** real people experience in the real world everyday. If you can find such a problem, and can adapt your idea to connect to it, your idea might become a real thing. If not, it is just a waste of your time to work on it.

Rather than calling it 'idea validation', 'validating problem-market fit' is a more accurate term. It sounds a bit abstract, but what 'problem-market fit' really means is that you're trying to find out if a certain problem resonates with your market: is there a sizeable portion of people that experience the problem? And do they care to have it solved?

To get to problem-market fit, you need to learn more about the problem.

In this stage, rather than looking at your undeveloped idea with a true/false mindset, you should go in discovery mode. It's about finding new information connected to the problem people experience. You need to fall in love with the problem and learn everything you can about it.

The experiments you'll run to support your search will be mostly **qualitative** and **explorative**. Only once you've found something interesting should you use **quantitative validation** experiments to see if your assumption is correct.

Forget about your solution

Perhaps your original idea had a specific solution in mind to solve that problem. In this stage, forget about that solution for now. It doesn't matter at this point *how* you are going to solve the problem. It matters if the problem *exists*.

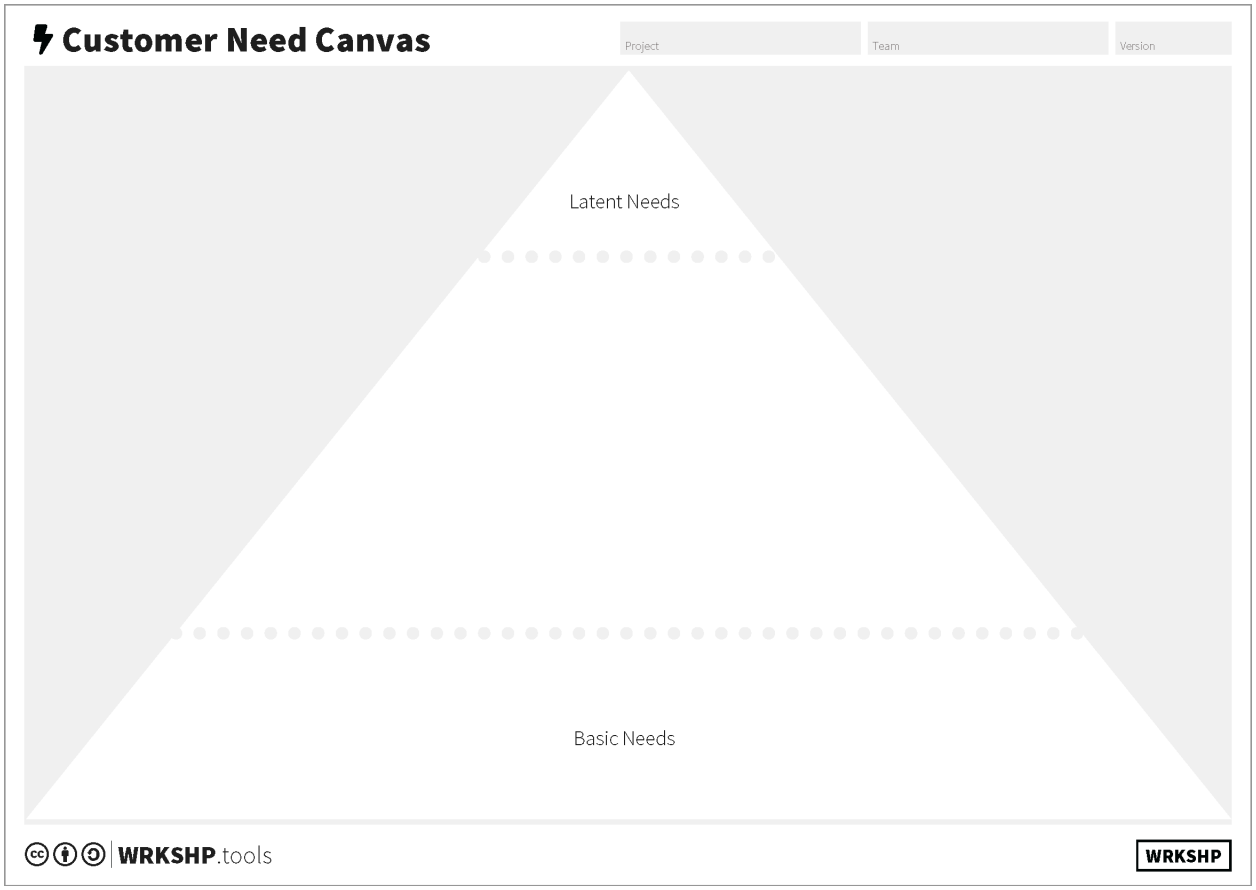
You also want to know what kind of problem you're looking at. If there isn't a true *need* for people to use your solution, if there isn't a *problem* your idea solves for them, then you won't overcome their natural inertia. People won't adopt your solution. They'll just continue living their lives as they did before.

- **Pains:** Problem that people experience every day, and that they can't avoid or get around except at high cost or effort.
- **Gains:** Needs that people may experience, but they can survive without it. Something that incrementally improves their quality of life.
- **Latent needs:** New possibilities that people don't know they need yet. An example: mobile phones were a latent need for most people in the 90s. Back then, most people simply could not imagine all the new use cases they use their phone non-stop for today. People simply don't know they need it yet.

Of these categories, pains are easier to turn into a product, and latent needs are the hardest. For a latent need, you'll probably have difficulty to find enough people that 'get it' early on to validate your results.

How to find out if you're dealing with a latent need

Think of the Maslow pyramid. At the wide base of the pyramid, you find things everyone needs, such as food, water, rest. These are things you can't do without. At the narrow top, you find things that only become important to people once their basic needs have been fulfilled.



If you think of problems to solve for your business, think about where that problem is in the pyramid. Is it on the bottom, or at the top? This is a way to visualize latent needs vs real needs.

The problem with latent needs is that people don't really know or care about fulfilling them yet. They are 'optional problems'.

That does not mean that a latent need can't be huge: in a 1999 interview, people in Amsterdam laughed at having a mobile phone in their pocket every hour of the day. They simply could not imagine what it would be like yet.

Interviews by Frans Bromet about the use of mobile phones in 1999 on the streets of Amsterdam. (English subtitles)

Convincing people of a latent need is an uphill battle. It can be done, but it is a lot harder than fixing something people are aware of as a problem.

The easiest way to tell if people have a real problem is to look for problems that people actually have a workaround for. Some way that they already solve that problem today. If that workaround

is unsatisfactory for them, for instance it is expensive, takes time, or only solves part of their problem, then you have a candidate for a real need.

[9.2] Preparing your Experiment

Where to find your test subjects

One of the most important aspects of running the experiment is where to find your test subjects. This very much depends on the kind of idea you have, and the problem you want to solve.

The key is to find people that are from the actual target audience — or what you think your target audience is at the moment — as possible. Friends and family are great to test your interview questions with, but they won't give you the information you need.

Think about the places (online and offline) where people experiencing the problem gather. What are the channels they use? What are the locations they visit? That's where you need to be to interview them.

Look back to section 5.2 in Preparing Experiments for more ideas on how to reach your test subjects.

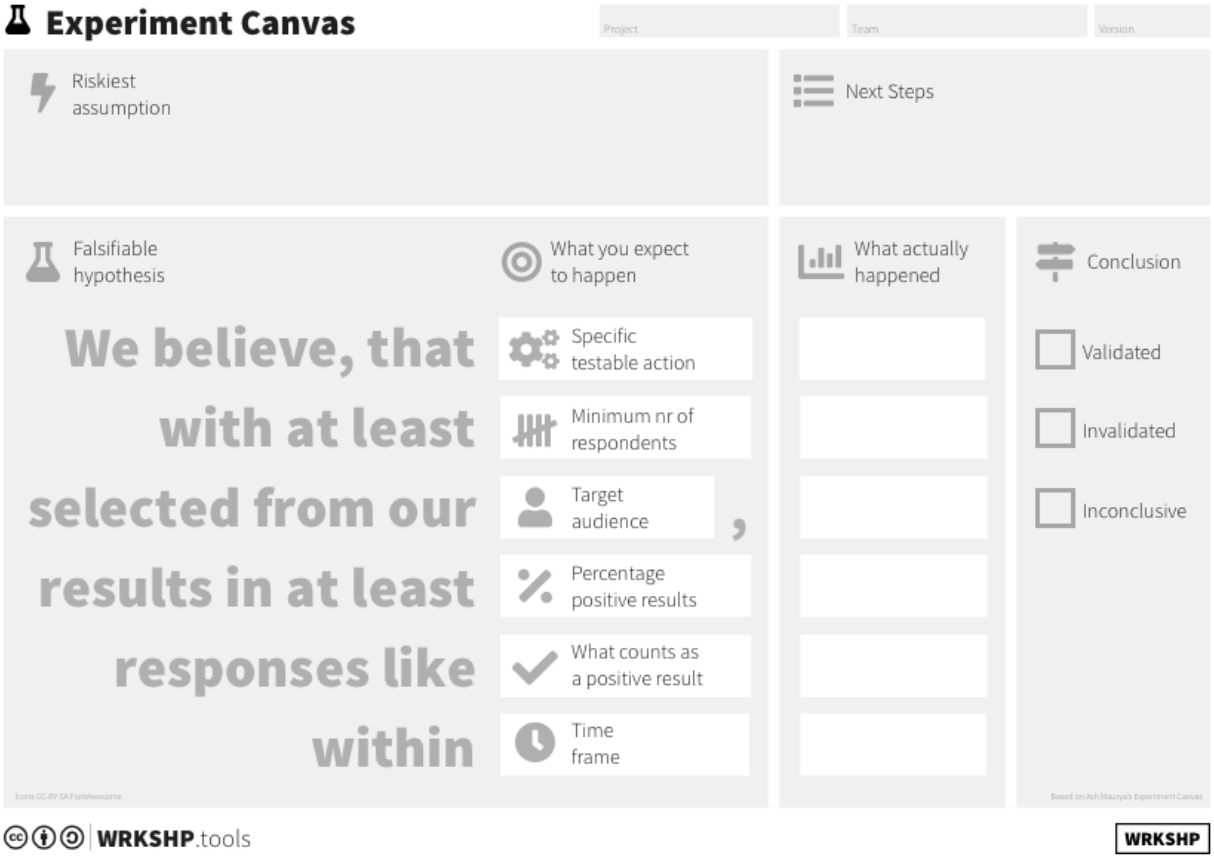
In this stage, you want to be very scrappy: chances are you'll be doing several of these explorative experiments, without being able to home in on a specific problem. Don't 'buy' or 'rent' test subjects, don't use advertising (unless it's free) - be creative and think of the best places, online and offline, to reach your test subjects.

If you already have access to a channel (e.g. an existing product, a well-visited website, a partner channel) with people from your target audience, all the better.

[9.3] Designing Experiments for Idea Validation

Because you're trying to uncover all kinds of qualitative information at this point, it is harder to run a clear-cut experiment where you test a hypothesis and based on the resulting data decide if the hypothesis was validated or not.

The information you get back will likely be vague, qualitative, and very much open to interpretation.



Use Ash Maurya's excellent experiment canvas to help define your experiment. See section 4.8 in designing experiments for more information on how to use this tool.

Still, defining risky assumptions and setting up an experiment with a clear hypothesis makes good sense. It helps you to focus what you're looking for. By defining how many people you want to talk to, and how many observations to make, and what you're looking for exactly in these observations, it is way easier to ward off confirmation bias. Just be prepared for a load of interpretable results. In this stage, a diversity of answers is good. (It's also not that big a problem if you're not seeing huge numbers of respondents, you'll need to double check what

comes out of this stage later with significant numbers. Better to do that when you know what to ask them.)

Recall module 4.3 and 4.4 in Designing Experiments and module 7.3 in Interpreting Experiment Results for extra input on how to best deal with qualitative information and explorative experiments.

Typical risky assumptions for idea validation:

- “People really have this problem”
- “People experience the problem on a daily basis”
- “People really care about this problem”
- “People would like to have a (better) solution”
- “It is a pain, not a latent need”

Typical hypotheses for idea validation:

- “We can find 10+ people that experience this problem through 50 interviews within 3 days.”
- “More than 70% of 20+ interviewed people say they experience this problem every day.”
- “More than 70% of 20+ interviewed people say the problem costs them over $100 per year.”

Example interview questions for idea validation:

- How many times did you experience the problem the last year?
- Can you tell me about the last time you experienced the problem?
- How did it make you feel?
- What happened or didn’t happen because of it?
- Do you have a workaround? Can you avoid the problem?
- What do you already do to prevent or avoid the problem from happening? Can you describe how you did that the last time it happened?
- How do you feel about the workaround?

Typical numbers of test subjects:

Experiments in this stage can be quite small, and typically range between 20-50 test subjects. Sure, that means statistically they are quite tricky. But as an entrepreneur, you need to balance getting a signal and spending too much time and effort. Talking to your potential customers will **always** give you valuable input, so it’s not a waste of time. Try to focus on big signals first and prove them with more specific experiments later.

[9.4] Interpreting Idea Validation Experiments

Interpreting the Results

When looking at the diverse and qualitative results you get out of explorative experiments, it can help a lot to differentiate them into categories:

- Quotes. What respondents actually said, and the types of words and descriptions they used.
- Perceived problem. How respondents described the problem and circumstances with their opinions, perceptions, and assumptions.
- Perceived needs. What respondents think they really need (what were they trying to achieve? What was their Job to be Done?). What other needs did they express? What solutions did they volunteer?
- Behaviour. The actual behaviours and activities respondents performed in the past that are related to the problem.
- Observations. What you as interviewer observed while interviewing.
- Conclusions. What in this collection of results impacts the idea (positively or negatively)
- Next steps. What other experiments can you run to get more information? What new ideas for (partial) solutions did you get? What further questions do you have about the problem?

The Experiment Result canvas is a great way to map out results with your team. Look at module 7.5 in interpreting experiment results for a step by step guide.

Besides using this method, make sure you capture the data in a clean, well organized sheet as outlined in 6.2, Running Experiments.

[9.5] Observing and Asking Questions

In the Idea Validation/Problem-Market fit stage, observing your test subjects' behaviour and asking great interview questions is super important.

It is very easy to ask biased questions and lead the test subjects to answers you'd like to hear.

Roy Fitzpatrick's excellent book 'The Mom Test' is a great read if you want to become an expert at asking questions. I won't repeat what he has written here, but I'll give some pointers nonetheless.

Don't ask for hypothetical situations
"Would you buy this product?" is a bad question. Why? Two reasons. First, people want to be nice and polite, and this question makes it very easy to give a socially acceptable answer. It doesn't *cost* the test subject anything to give an untruthful answer. Second, even if the test subject has no social skills and just gives an 'honest' answer, people can't predict how they will behave in a hypothetical situation with any accuracy. So the test subject might still lie, inadvertently.

Don't ask for their opinion of your idea
Ask about their life, their experiences, and their way of dealing with a problem. It's your job to connect that information to your idea, but you best do that afterward. The test subject doesn't need to know or like your idea. The experiment is not made to inflate your ego or remove that feeling of uncertainty in your stomach. It's about data.

> You aren't allowed to tell them what their problem is, and they aren't allowed to tell you what to build.
>
> -- Rob Fitzpatrick

Ask for behaviour
Instead of a hypothetical situation, ask what test subject's behaviour was for real situations they experienced in the past. When was the last time they experienced this particular problem? What was it like? What did they do about it? In the end, it's not what people say that is important, it's what they do.

Ask for commitment
Make it more difficult for test subjects to (inadvertently) 'lie' to you, by asking them to make a commitment. The best way to know if a customer wants to buy a product is to try and sell it to them. The second best way is to ask for some other token of their interest: their email address,

access to their network, their time, their reputation, you name it. Asking someone if they will publicly tweet about your product under their own name will give you a much more honest answer.

Ask the important questions
It's hard to ask the questions where a customer may tell you they hate your product. It's much easier to stay away from those questions and ask irrelevant questions. This way, your ego won't get heard. But by avoiding these 'hard truths', you will also not learn what you really have to fix in your product.

> You should be terrified of at least one of the questions you're asking in every conversation.
>
> -- Rob Fitzpatrick

Don't sell
You're trying to get information from your customer, you're not trying to convince them. It's hard to not be enthusiastic about your idea, I get it. But by selling and convincing the test subject you're not getting the data you need. You should be the one to be quiet and listen, and make the test subject do the talking.

When you introduce the idea, it's okay to do it indirectly. You don't need to introduce it as your own idea. You're just curious what they think about it. By removing yourself from the equation, you'll avoid a lot of politeness and socially acceptable answers.

And finally, remember that **anything the test subjects say is right and valuable**. It's not your job to correct them or make them see the light. There are no 'wrong answers'. There will be plenty of time for selling later.

Look for the right anwers
When you interview people and (eventually) show them your idea or prototype, you will get different types of results. Which of the below results is a good sign? And which is a bad sign? To know that, you should look for the willingness of the test subject to make a commitment.

"That's so cool! Love it!"
Bad result. There is no commitment.

"Looks great! Let me know when it launches!"
Fail. No commitment.

"I would definitely buy that"

Fail.

“When you’re ready, I can find some people you can show it to.”
Mostly fail.

“What are the next steps?”
Success.

“Can I buy the prototype?”
Success.

Example interview questions for idea validation:

- How many times did you experience the problem the last year?
- Can you tell me about the last time you experienced the problem?
- Please describe how it made you feel
- What happened or didn’t happen because of this problem?
- Do you have a workaround? Can you avoid the problem?
- What do you already do to prevent or avoid the problem from happening? Can you describe how you did that the last time it happened?
- What does this problem cost you?
- How do you feel about the workaround?

[9.6] Exercise: Good and Bad questions

In this exercise, you see a list of questions that you might ask in interviews. Decide if this is a good or a bad question to ask, and why. If it's a bad question, how might you change it so it becomes a good question?

1. Imagine your company failed. Why did it fail?
2. Would you buy this product?
3. How much should this product cost?
4. What was the last time you have experienced problem X?
5. Do you think your manager will find this interesting?
6. Have you ever bought an app? What did you pay?
7. Do you think solution Y solves your problem?
8. Why did you choose solution Y? What were your alternatives?
9. Did you try alternative solutions?
10. What would happen if you did not solve problem X?
11. Can you introduce me to your manager so I can show them this prototype?
12. Can I have your email address so I can keep you informed of future developments?
13. We have a demonstration of the prototype next wednesday. It would be great if you could come to see it.
14. Would you refer three people from your network for me to talk to about this subject?

Make up three more 'good' questions yourself.

[9.7] Recipe 1. Friends & Family

About the experiment

The ‘Friends and Family’ experiment is a simple, fast experiment to get some useful feedback on your idea and to learn how other people experience the problem you’re interested in. The term ‘friends and family’ should be seen loosely: you’re interviewing people you already know well. Close business relations and well-meaning colleagues are definitely potential test subjects for this experiment.

This type of experiment is usually explorative. It’s more like an open conversation than a strict questionnaire. The conversation may be steered in new directions that you were not aware of before. That’s ok. Because of the small size of the pool of test subjects and the way they were selected this experiment is unsuitable to base important decisions on, but is very valuable as a first step and a dry run for other experiments.

Datasheet

Experiment: Friends & Family
Type of experiment: Explorative, qualitative.
Benefits:

- Easy to set up
- Fast
- You already have access to the test subjects

Dangers:

- Friends and family will most likely (inadvertently) ‘lie’ (Mom Test) because they have a vested interest.
- Self-selected audience which probably does not adequately reflect the target audience.

Use this experiment to: Get the first feedback for your idea. The friendly audience will mean that you avoid the risk of failure and rejection, which can smooth the road ahead. If your friends and family don’t like it or don’t understand it, then nobody will. You can also use this as a dry run for a larger interview.
Time and resources required to set up: A Friends and Family experiment can be designed and run in a day or at most a few days, and doesn’t need to cost anything besides time.
Evidence level: Weak. Your audience selection is probably not adequate, and you’ll be exploring, so you’ll get qualitative results that are hard to score or align. There may be interesting patterns in the data.
Method: The easiest way is to do this as a face-to-face or telephone interview. (Phone interviews are hard with people you don’t know, but with people you have met and know well it’s easier)

Prototype: Interview questions

Step by Step

Step 1. What are you going to explore?

The first thing to decide is if you're going to create an explorative experiment or a validation experiment. If you're still quite fresh to the problem space and you are trying to learn how other people think about the problem, start with an explorative experiment.

This experiment, because you probably won't have a friends and family network of 100s of people, is typically explorative and is best used as a first step towards understanding how other people see the problem.

The first thing to do is to scope out what (you think) the problem is you want to learn about. When do you think it manifests? How do you expect people to experience it? This is your baseline. Your riskiest assumption will be that people indeed experience the problem you have defined.

Based on this, try to fill out the hypothesis. Since it's not a validation but an exploration experiment, it's okay that the numbers are a bit fuzzy. Still, you need to make some sort of prediction about your friends and family: how many will answer positive?

With your team, find out how many test subjects you can realistically approach.

Step 2. Prepare

Make a list of names to interview per team member, with their contact details. Set a time limit (this experiment is best done fast, and typically does not need to take more than 1-2 days time). Set aside time to do the interviews.

Have each team member contact their prospective test subjects as soon as possible to plan a time to interview them. The sooner you find out people aren't available, the sooner you can find replacements. If it turns out to be difficult to talk to everyone in the time frame you set, you can talk about extending it.

Create a Google Sheet for you and your team to fill in.

	Subject A	**Subject B**	**Subject C**
Question 1	Answer A.1	Answer B.1	Answer C.1
Question 2	Answer A.2	Answer B.2	Answer C.2

Tip: *Create different tabs in Google Sheets for each team member so that it's clear where to put their data.*

Step 3. Create a script with questions

The next step is to create a list of questions. Refer back to section 8.6 (Observing and asking questions) for inspiration.

Remember you want to explore if and how your test subjects experience the problem. The questions you ask should be geared towards having them talk about that situation. Don't immediately tell them about the problem, but rather bring up the circumstances you're interested in. List the questions in the sheet.

You'll probably find out during your interviews that your audience doesn't immediately connect in the way you'd like. That's ok. This is exactly why this experiment is such a good 'dry run' for a larger interview experiment. With your friendly audience you'll have a second or even third chance to explain what you would like to know, and they'll help you find a clear direction.

For each question, also note down what you'll be looking for in the answer to make it count as positive.

Example: A startup had the idea to build an app to help small business owners get real-time insight in their financial situation. I was brought on board to help them move away from a solution-centered mindset and towards customer validation. The first thing we did was run a 'friends & family' experiment, with friendly business relations and partners in the target audience. We wanted to understand if small business owners experienced the problem.

Rather than ask them straight away, we started with a few onboarding questions about how their business was doing. Was it doing well, were there exciting new developments. Most business owners became quite engaged when talking about their business. Then, we steered the conversation towards finance, with questions such as:

- Are you completely up-to-date financially with your business?
- How often do you take time to get insight in your finance?
- How much time do you spend going over finances?
- What do you typically look for?
- How do you manage your finances today?
- Are you happy about how that is going?

We quickly found out that most of the more established SME's we talked to had their finances covered. They used an accountant, a savvy employee, or a stack of tailored tools and spreadsheets to have a grip on their finances. We learned that the established firms were

probably not the right audience, but that startups, new businesses, and especially freelancers were struggling, reinventing the wheel, and making costly mistakes.

We learned this from doing 20 interviews in 3 days, and got very valuable information out of it that required some serious rethinking.

This shows you that the first conversations with (potential) customers can be extremely valuable for the rest of your journey. There really is no excuse not to spend 2-3 days talking to people. The worst thing that can happen is that your idea doesn't resonate and they prevent you from making a costly mistake.

Step 4. Run the experiment

Have team members run their interviews and keep notes in the Google Sheet.

Tip: *You'll be able to see how many interviews have been done real-time, and if you see team members lag behind, contact them to find out what's wrong. In real-life situations, it often happens that people overcommit and underestimate the time requirements of interviews. It's better to extend the deadline a bit and give people the time they need to do interviews when you see this happening. Setting a long deadline beforehand will only be counterproductive since there is no sense of urgency.*

Step 5. Interpret the data

Use the Experiment Outcome canvas (see section 7.5) to discuss your results with the team. Use the scoring method you created in step 3 to go over the results and see if you can find any interesting patterns in the data.

Are there clear conclusions to be drawn? What are signals that you find interesting and want to learn more about?

[9.8] Recipe 2: Exit Polls

About the experiment

The ‘Exit Poll’ experiment aims to get a reaction from test subjects just at the time they have experienced the problem (or at least, that’s what you assume). If you want to know how people experience their commute, ask them when they come out of the train station, or park their car. If you want to know what people think about when they are purchasing a vacuum cleaner, ask them when they just bought one.

Most often, an exit poll experiment takes the shape of an interview. You’ll find a suitable location frequented by people you think experience the problem, and try to approach people and convince them to do a quick interview. In the interview, you’re most interested in the experience they just had, and what they think about that.

Datasheet

Experiment: Exit Poll
Type of experiment: Mixed (exploration or validation)
Benefits:

- Easy to set up
- Fast
- Possibly better in line with your target audience

Dangers:

- You need a good location to reach test subjects
- Can be hard to get test subjects’ attention

Use this experiment to: Quickly get feedback from a larger group of people about how they experience a problem, get first feedback if people are interested in your idea, get some qualitative data
Time and resources required to set up: You’ll need a bit of time and manpower to pull this off. You’ll probably need a day to set it up, and a day to run it.
Evidence level: Weak. You’ll get information, but you won’t easily get a commitment from people you talk to on the street. People self-select: they determine they want to spend time talking to you, which can skew results.
Method: A good way to do it is to approach people in a busy location and ask them questions (or have them take action, see the hack below)
Prototype: Interview questions, something to draw their attention

Step by Step

Step 1. What do you want to explore vs your riskiest assumption

If you're approaching this as an exploration experiment, find out what you want to know about. What are your assumptions about how people experience the problem? What do you expect they will tell you?

If you're approaching it to validate, you need to pinpoint your riskiest assumption. Check the default riskiest assumptions for the stage you're in first, and use the Riskiest Assumption canvas to make sure you have the right one.

Next, build your hypothesis. Find out how many people you think you can approach reasonably, and how many you predict will answer positively.

Tip: To get an idea of the amount of people you can reach, use your calculator. Take the amount of time it takes to get through your questions (typically this should be 2-3 minutes max for this experiment), and multiply that by four. You'll need time to convince your next subject to be interviewed. Take the time you want to spend in the location, and divide that by the time per interview. Multiply by the number of teams, and that's your upper bound.

Example: We did an 'exit poll' experiment at WebSummit 2018, where we talked to startup founders. (Read more about this experiment here) We calculated as follows:

Teams: 2
Estimated time per interview: 5 minutes
Time to find a test subject: 10 minutes
Time on venue: 4 hours per day
Days on venue: 4

Total interviews per day max: 4 * 2 * (60/15) = 32 max.
Total interviews for four days: 128 max.

In the end, we managed around 50 in total. We underestimated how much doing so many interviews would wear us down, and successful interviews sometimes were more than 15 minutes long. Also, because we were only two people, each of us had to interview and write down the answers at the same time, which was impossible to do. Between interviews, writing down the information took time as well.

Step 2. Prepare

If at all possible, visit the venue at a time slot similar to when you want to do your interviews. How many people are there? What are they doing? Are they in a hurry? And, also quite invaluable, do you have any competition?

It's tough to get people to stop and talk to you when you're in a spot where passers-by are used to being harassed by people flyering, selling, or doing other surveys. If you see other people do this, think about getting another location, or think about a way to signal that you are not selling anything when you do your experiment.

Best thing to do in preparation is to try to approach someone and note how they respond.

Create a Google Sheet for you and your team to fill in, same as before.

	Subject A	**Subject B**	**Subject C**
Question 1	Answer A.1	Answer B.1	Answer C.1
Question 2	Answer A.2	Answer B.2	Answer C.2

Tip: *If you're going to count certain behaviours or answers, it may be easier to print out your sheet and keep a tally, or to use a tally counter on your phone, such as Tally 2. This can make it much easier to keep track of how many times different answers or actions occurred.*

Step 3. Create a script with questions

The next step is to come up with suitable questions. Focus on questions about their recent experience. Add the questions in the sheet. Also, for each question, think about what answers you'll be looking for to validate your assumption. What counts as a 'positive' answer for you?

Example: Questions asked to public transport commuters at the train station

- How was your journey today?
- When did you start your journey today?
- How often do you travel by train?
- How does your experience compare to other times?
- What did you love / hate about it?
- Would you recommend going by train to others? Why (not)?

It can be difficult to approach people and talk to them. It is socially awkward. It can help to fabricate an 'excuse' to make approaching people easier. Anything that draws (positive) attention can work. Think of a good smile, brightly colored t-shirts, or something else that makes

people curious. Giving away something for free can work wonders as well. If you can come up with something that draws a crowd of people, you're going to have an easier task to interview them.

When you come up with something, it can help to make it look not too shiny and well organized: people will be suspicious you're trying to sell them something. Make it look fun and friendly.

Just don't stand around with a clipboard and fear in your eyes, that's a sure way of making people hurry past you.

Step 4. Run the experiment

When it is time to run the experiment, brief the team well. Make sure you arrive at the location early, and that you bring enough refreshments and drinks to keep everyone going. When possible, use a car parked nearby or a table at a coffee place as a 'home base', where interviewers can take some rest and discuss recent findings. Keep track of how fast interviews are coming in, and take action if it is not going according to plan.

Tip: *It can be difficult to get started. If you're not doing this every day, especially the first few approaches will be difficult. Tell everyone this is expected, and promise an award for the person that first breaks the ice. Framing it as a game reduces a lot of the tension.*

Step 5. Interpret the data

Use the Experiment Outcome canvas and your sheet to go over the results. Was it easy to approach people? How did team members experience it? What were the most surprising results?

First, check if you have enough data. Did you meet the quota you set in the hypothesis?

Next, score the answers and see if your score validated the assumption. Because this typically still is a low-n experiment, take the outcome with a grain of salt if your signal was so-so.

If you find the experiment did not get a clear outcome or if you did not have enough data, think about extending the experiment or running it in a different location or time slot. What could you do differently to make your signal clearer? What was the single question that gave you the most information?

Hacks

A great hack that turns this experiment into a qualitative experiment, and makes it much easier to get people to contribute, is to set up two 'gates' passersby can select to walk through. Put them a little out of the normal route, so people have to go through them deliberately. Mark each with a clear one-line statement, and count people going through the gates.

We did this to find out what people thought of their commute with the Dutch Railway, at the exit of the train station. We set up three traffic cones, and market one gate with “I love my commute” and the other with “I hate my commute”. About 50 people deliberately walked through the gates in half an hour, most of whom said they hated their commute that day. We asked a few of the test subjects follow up questions, if they were open to that.

Warning: *this experiment ended for us because we got kicked out of the train station. Make sure you find a spot where you don’t get expelled, or ask permission :)*

Another hack a team in an innovation workshop came up with is to use an object to make people curious. The subject the team wanted to learn about was how people transport groceries to their urban home. They came up with a (scrappy) prototype of a robot that can carry your groceries for you. We ran the workshop in-house in a large corporate, and they decided to try and stop people going to the in-house cafe for lunch to ask them questions. People definitely were curious about this cardboard-and-plastic robot prototype, but when they were on their way to have lunch they didn’t stop to talk. On their way back, however, it was much easier. They obviously had been thinking about this out-of-the-ordinary sight over lunch and wanted to know more.

Case study: Main Ingredient Web Summit post

[9.9] Recipe 3: Map the Problem

About this experiment

Map the problem is another research experiment. You'll use information from different sources (including potential customers) to create a 'map' of the problem. This will help you understand the problem and your customers better.

Basically, you are trying to piece together everything that you can find related to the problem into one comprehensive big picture. During your innovation journey, you'll be able to refer back to this map and update it with new information.

Datasheet

Experiment: Map the Problem
Type of experiment: Explorative, qualitative
Benefits:

- Get an in depth view of the problem and everything related to it
- Understand the problem better

Dangers:

- Can be a bit of work to get right

Use this experiment to: Get a complete picture of your problem
Time and resources required to set up: You need a big sheet of paper and time to do research and customer interviews
Evidence level: N.A, it will give you new ideas and elicit risky assumptions however
Method: Do research and interviews to map out the problem
Prototype: N.A.

Step by Step

Step 1. List your questions

What aspects of the problem do (you think) you know? And what areas do you feel out of synch with? You should be able to answer the list of questions below to complete your map, but you may need to dive deeper.

- What does the problem cost? (And to whom?)
- What can people not do because of this problem?
- What do people have to do because of this problem?

- How often does it happen?
- Why does it happen?
- What are causes or contributors?
- To whom does it happen?
- Who are the ‘biggest victims’?
- What are current solutions and workarounds?
- What do people like/dislike about these solutions?
- What kind of problem is it (nice to have, need to have)?
- Is there a group of people for whom it’s a need to have?
- Is there a group that is willing to pay to solve it?
- What are they currently paying for (alternative) solutions?
- Is there anyone who benefits from this problem?
- …

Step 2. Run the experiment

Try to find out as much as you can using desk research first, and then go and talk to potential customers, people who experience the problem, and other stakeholders. Try to check your data with them, and ask for more input.

Step 3. Analyze your data

Once you have compiled a lot of data, go over it. Can you organize it or structure it in a meaningful way? Are there different categories or scales of the problem, different groups of people that experience it? Did you find out a cause or key drivers or contributors?

- Try to plot your map on a timeline. Where do different stakeholders show up?
- Try to plot your problem in the center and different customer segments around it. How does the problem affect them?
- How are all the different parts connected?
- What is the problem’s achilles heel?

Sources							
Reddit	Search for 'experiments' on Reddit r/startups						
Reddit	Eric Ries AMA on Reddit						
Quora	Search for 'lean startup' on Quora						
...							
Posts							
Source	**Title**	Link	Post	**Question**	**Keywords**	**Answers**	**Notes**
Reddit	Building a landing page to test product market fit.	https://www.reddit.com/r/startups/comments/bqpkm6/building_a_landing_page_to_test_product_market_fit/	We have gotten to the point of our project where we want to build a landing page for our b2b web based business to see how many users sign up. Over the past few months I have gotten a lot of good verbal responses from potential users that they would use our service but, before we start spending resources on building the entire site we want to be sure. The question that I have is building a landing page for a b2b web based platform and gathering sign ups a good way to gauge interest in the business? One side of the network is going to be suppliers and the other side is going to be consumers. We need to have both sides use the service for it to work and we believe we are adding enough value to both sides to do so. From a marketing standpoint for the landing page I was going to use Facebook marketing and also LinkedIn to directly target the users we seek. Our web developer said the landing page would cost around 1000$ and I was going to spend another 1000$ on marketing at first to see where that leads us and then reevaluate from there. If we can get our site in front of the eyeballs of our potential target customers, we believe the conversion rate to sign up will be high as it's a free service for both sides.	Is building a landing page for a b2b web based platform and gathering sign ups a good way to gauge interest in the business?	Landing page, Verbal responses, making sure, gauge interest, marketing, target customers, cost, see where that leads, re-evaluate, conversion rate	Use something like instapage. You can make landing pages yourself and test different copy and layouts. Don't spend so much on developing a single landing page. | $1000 landing page is insane, ive worked with start ups using god awful unsplash squarespace websites, as long as marketing and UX/UI is okay you should be able to get by. | Your expenses on a landing page experiment should be pretty close to $0 for up to 2 weeks and maybe $200 for a few months. |	Nobody answered the actual question
Reddit	How do you keep track of product experiments?	https://www.reddit.com/r/startups/comments/bjhob7/how_do_you_keep_track_of_product_experiments/	Good morning product owners! I have a small product that's starting to gain a little bit of traction. I've defined a few key metrics based on the pirate metrics framework. I'm currently focusing on engagement and iterating through product versions which introduces subtle changes to test out how it impacts that metric. I have 2 questions in regards to product experimentation. Are you applying some sort of formal process to test out ideas on your product? And how do you keep track of those experiments, such as the metric you want to influence, the impact of the experiment, etc?	Are you applying some sort of formal process to test out ideas on your product? And how do you keep track of those experiments, such as the metric you want to influence, the impact of the experiment, etc?	metrics, traction, product owners, pirate metrics, emgagement, iterating, formal process	I'm using airtable to track everything. But that's just a tool, you should use the one you like the most... But I'm seeing that you are missing the methodology more, maybe? | I maintain an dropbox paper document that lists my initial hypothesis, metric that I want to improve, what could go wrong , side effects of the experiment (any metric that went down because of the experiment) and learnings from the experiments. I also maintain another master document that lists all my experiments and their learnings. |	Focus is on tools and technology in the answers
Reddit	How do you keep track of product experiments?	https://www.reddit.com/r/startups/comments/bjhob7/how_do_you_keep_track_of_product_experiments/	I'm also having the trouble of tracking my experiments -- the changes I made to the copy/product -- and how it affect my core metrics.	How to track my experiment? How do changes to the product change my core metrics?	tracking, metrics	Well, my process is far from perfect. I currently document my experiments in Asana with a bunch of custom fields. Those are exported to my data warehouse (Snowflake), I do some transformation of the data (with dbt) and analyse it with R.	
Reddit	How do you keep track of product experiments?	https://www.reddit.com/r/startups/comments/bjhob7/how_do_you_keep_track_of_product_experiments/	How do you document your experiment? I might be able to learn a thing or two because I tried writing down the experiments, the numbers, the conclusion and all that stuff on evernote but it gets quite messy.	How to document your experiment?	document, experiment, numbers		
Reddit	Product Market-fit experiment. Part of the main product or independent?	https://www.reddit.com/r/startups/comments/bt3sgu/product_marketfit_experiment_part_of_the_main/	Hello, I am the CEO of the fitness tracking platform. At the moment I have an idea which I want to validate for the same customer segment. It might be a part of the main product in the future. I am thinking if I should create an MVP separately to validate how people react or make it a part of the main product to continue acquiring users. From my understanding, if I make it as a separate experiment it might help me to a valdiate better idea. What is your advice?	Should I create a separate MVP for a new customer segment?	MVP, validate, customer segment, experiment	Yeah lots of companies use A/B testing or Triangle on one side Circle on the other ... You could try and segment the population so approx half your users get the feature change. You need to add some sort of feedback mechanism that's effortless, too. I think you could do it, but A/B testing is not necessarily the best way to take creative control of an app. If you're confident in your move, just do it, A/B testing is typically for honing and dialing in things way later down the line. You start getting blessed by statistics instead of the muses.	Answer is again technological

Listing:

- The source
- Title of the question or post
- Link
- The post text (or relevant pieces)
- Some relevant answers

When taking a day or so to do this mind-numbing work, you'll be able to start tallying things like keywords, and seeing patterns in the types of questions.

Doing this immediately starts to give you new ideas and insights. There are a lot of questions about:

- Unclear definitions
- When to use a certain Lean Startup technique
- What to do in a specific situation
- What experiment to use and how to use it
- Variations on 'does Lean Startup really work?'

The last type of question is interesting, because it shows that the preconceptions about lean startup, 'testing EVERYTHING', are misguided. The concept of testing the right thing is apparently not clear to everyone.

Doing research like this can help you define what people may be struggling with. Always do remember that any online forum is a self-selected space. Certain types of people like to use these forums, and others don't. So the questions you find may be different from the questions your target audience as a whole has.

It's also interesting to make a few observations of your own. Some posters seem to be looking for certainty or security before committing to an experiment. They want to check if they're setting it up right, if their assumptions make sense, etc. The answer usually is 'run the experiment and find out'. I wonder if that addresses the need the poster had when they posted the message.

Hacks

A hack you can do with these experiments is to try to contact the posters of interesting questions, and ask them if you can interview them. It can sometimes be done through a direct message, or perhaps you can find the person on LinkedIn. If you are lucky, you can get much deeper information in this way, including how they solved their problem and if they are satisfied with that solution.

[9.10] Recipe 4: Online Research

About the Experiment

To understand your target market better and figure out if and how they experience the problem you want to solve for them, a shortcut is to use online research.

This experiment recipe is something you can even do when you have not a single customer. It's a variant of market research, and uses existing online forums like Reddit, Quora, and Facebook groups to extract relevant questions and answers.

Rather than going out and painstakingly asking questions yourself, you're tapping into the collective questions of hundreds of thousands of people in your target audience that are using these forums every day. And you know that when they make the effort to post a question they certainly care about that subject.

How many people ask questions about this topic? How many people review products? What communities exist? What are they talking about?

It's time to get online and start tallying and analyzing questions.

Datasheet

Experiment: Online Research
Type of experiment: Explorative, qualitative.
Benefits:

- Easy to set up
- Relatively fast
- You get insight in the most asked questions
- You should be able to find good sources for most assumptions

Dangers:

- It can be time consuming to go over everything in detail
- You get a lot of data that may not be a 1:1 match for your purposes
- Self-selected audience which probably does not adequately reflect the target audience.

Use this experiment to: Get an idea of the problems your target audience struggles with, and how they solve these problems. Learn what questions they ask, and how they think.
Time and resources required to set up: You'll probably need to become a Reddit or Quora member, and it will take some time.
Evidence level: Weak to moderate. People are not guided by your bias in the questions they ask, and it is possible to see how often some questions appear or how popular they are.

Method: A laptop and a google sheet to record findings.
Prototype: None

Step by Step

Step 1. What are you going to explore?

The first thing you need to do is to figure out what you want to know. If you already have assumptions, use the riskiest assumption canvas to find the most important ones. Otherwise, try to come up with some questions you think your audience is asking. Define a clear hypothesis to test.

Define a list of keywords that fit with your questions.

Tip: *if you're in a very jargon-prone market, remember that your potential customers may use different words to describe things! They may be searching for 'how to build my website' rather than 'Full CMS page builder' for instance.*

Step 2. Find sources of online data

Make a list of forums and other sources people from your target audience may be using. Think of things like:

- Reddit and Quora
- LinkedIn groups
- Facebook groups
- Product review sites
- Competitors' facebook or forum pages
- Industry or market specific forums (***Tip:*** *ask a customer what sources they use*)
- Medium posts
- …

There are tons of places you can find for your situation. Spend a half-hour or so to come up with a first list (and save it, you'll be adding more sites to it later).

Step 3. Look for relevant posts

Use the on-site search of each source to search for posts and questions based on your keyword list. You'll get a list of results.

Go over the answers one by one, and copy relevant posts and answers in your sheet. Make sure save the link to each post as well.

Step 4. Analyze for patterns

Go over the posts you saved. Are they in line with what you expected? Or are they very different? What jumps out at you?

Next, try to find a pattern in the types of questions. Categorize them in practical or theoretical answers, try to tag them with keywords relevant to your business. Are there questions that come back multiple times?

You can use the Experiment Result canvas to map out responses with your team and discuss what is in line with expectations, and what is not.

Make a shortlist of things that were new to you and that you want to know more about, or that you have to add to your list of assumptions.

Example:

I did this experiment as well, when I was building this course. Below you can see some of the materials I uncovered.

Exploration Question: What are people struggling with when they do lean innovation and experiments?

Specifically:

- Are people asking questions about Lean Startup?
- Do people need help selecting experiments?
- Do people need help with doing experiments?

Sources (found using google):
Search for ‘experiments’ on Reddit r/startups
Eric Ries AMA on Reddit
Search for ‘lean startup’ on Quora
…

A number of interesting statements and questions pop up. I made a google sheet doing research that looks like this:

A number of interesting statements and questions pop up. I made a google sheet doing research with the following columns

- The source
- Title of the question or post

- Link
- The post text (or relevant pieces)
- Some relevant answers
- Tags for each post
- My notes

Sources							
Reddit	Search for 'experiments' on Reddit r/startups						
Reddit	Eric Ries AMA on Reddit						
Quora	Search for 'lean startup' on Quora						
...							
Posts							
Source	**Title**	Link	Post	**Question**	**Keywords**	**Answers**	**Notes**
Reddit	Building a landing page to test product market fit.	https://www.reddit.com/r/startups/comments/bqpkm6/building_a_landing_page_to_test_product_market_fit/	We have gotten to the point of our project where we want to build a landing page for our b2b web based business to see how many users sign up. Over the past few months I have gotten a lot of good verbal responses from potential users that they would use our service but, before we start spending resources on building the entire site we want to be sure. The question that I have is building a landing page for a b2b web based platform and gathering sign ups a good way to gauge interest in the business? One side of the network is going to be suppliers and the other side is going to be consumers. We need to have both sides use the service for it to work and we believe we are adding enough value to both sides to do so. From a marketing standpoint for the landing page I was going to use Facebook marketing and also LinkedIn to directly target the users we seek. Our web developer said the landing page would cost around 1000$ and I was going to spend another 1000$ on marketing at first to see where that leads us and then reevaluate from there. If we can get our site in front of the eyeballs of our potential target customers, we believe the conversion rate to sign up will be high as it's a free service for both sides.	Is building a landing page for a b2b web based platform and gathering sign ups a good way to gauge interest in the business?	Landing page, Verbal responses, making sure, gauge interest, marketing, target customers, cost, see where that leads, re-evaluate, conversion rate	Use something like instapage. You can make landing pages yourself and test different copy and layouts. Don't spend so much on developing a single landing page. | $1000 landing page is insane, ive worked with start ups using god awful unsplash squarespace websites, as long as marketing and UX/UI is okay you should be able to get by. | Your expenses on a landing page experiment should be pretty close to $0 for up to 2 weeks and maybe $200 for a few months. |	Nobody answered the actual question
Reddit	How do you keep track of product experiments?	https://www.reddit.com/r/startups/comments/bjhob7/how_do_you_keep_track_of_product_experiments/	Good morning product owners! I have a small product that's starting to gain a little bit of traction. I've defined a few key metrics based on the pirate metrics framework. I'm currently focusing on engagement and iterating through product versions which introduces subtle changes to test out how it impacts that metric. I have 2 questions in regards to product experimentation. Are you applying some sort of formal process to test out ideas on your product? And how do you keep track of those experiments, such as the metric you want to influence, the impact of the experiment, etc?	Are you applying some sort of formal process to test out ideas on your product? And how do you keep track of those experiments, such as the metric you want to influence, the impact of the experiment, etc?	metrics, traction, product owners, pirate metrics, emgagement, iterating, formal process	I'm using airtable to track everything. But that's just a tool, you should use the one you like the most... But I'm seeing that you are missing the methodology more, maybe? | I maintain an dropbox paper document that lists my initial hypothesis, metric that I want to improve, what could go wrong , side effects of the experiment (any metric that went down because of the experiment) and learnings from the experiments. I also maintain another master document that lists all my experiments and their learnings. |	Focus is on tools and technology in the answers
Reddit	How do you keep track of product experiments?	https://www.reddit.com/r/startups/comments/bjhob7/how_do_you_keep_track_of_product_experiments/	I'm also having the trouble of tracking my experiments -- the changes I made to the copy/product -- and how it affect my core metrics.	How to track my experiment? How do changes to the product change my core metrics?	tracking, metrics	Well, my process is far from perfect. I currently document my experiments in Asana with a bunch of custom fields. Those are exported to my data warehouse (Snowflake), I do some transformation of the data (with dbt) and analyse it with R.	
Reddit	How do you keep track of product experiments?	https://www.reddit.com/r/startups/comments/bjhob7/how_do_you_keep_track_of_product_experiments/	How do you document your experiment? I might be able to learn a thing or two because I tried writing down the experiments, the numbers, the conclusion and all that stuff on evernote but it gets quite messy.	How to document your experiment?	document, experiment, numbers		
Reddit	Product Market-fit experiment. Part of the main product or independent?	https://www.reddit.com/r/startups/comments/bt3sgu/product_marketfit_experiment_part_of_the_main/	Hello, I am the CEO of the fitness tracking platform. At the moment I have an idea which I want to validate for the same customer segment. It might be a part of the main product in the future. I am thinking if I should create an MVP separately to validate how people react or make it a part of the main product to continue acquiring users. From my understanding, if I make it as a separate experiment it might help me to a valdiate better idea. What is your advice?	Should I create a separate MVP for a new customer segment?	MVP, validate, customer segment, experiment	Yeah lots of companies use A/B testing or Triangle on one side Circle on the other ... You could try and segment the population so approx half your users get the feature change. You need to add some sort of feedback mechanism that's effortless, too. I think you could do it, but A/B testing is not necessarily the best way to take creative control of an app. If you're confident in your move, just do it, A/B testing is typically for honing and dialing in things way later down the line. You start getting blessed by statistics instead of the muses.	Answer is again technological

You can find an example sheet here

When taking a day or so to do this (mind-numbing) work, you'll be able to start tallying things like keywords, and seeing patterns in the types of questions.

Doing this immediately starts to give you new ideas and insights. There are a lot of questions about:

- Unclear definitions
- When to use a certain Lean Startup technique
- What to do in a specific situation
- What experiment to use and how to use it

- Variations on ‘does Lean Startup really work?’

The last type of question was interesting to me, because it shows that the preconceptions about lean startup, ‘testing EVERYTHING’, are misguided. The concept of testing the right thing is apparently not clear to everyone.

Doing research like this can help you define what people may be struggling with. Always do remember that any online forum is a self-selected space. Certain types of people like to use these forums, and others don’t. So the questions you find may be different from the questions your target audience as a whole has.

It’s also interesting to make a few observations of your own. Some posters seem to be looking for certainty or security before committing to an experiment. They want to check if they’re setting it up right, if their assumptions make sense, etc. The answer usually is ‘run the experiment and find out’. I wonder if that addresses the need the poster had when they posted the message.

Hacks

A hack you can do with these experiments is to try to contact the posters of interesting questions, and ask them if you can interview them. It can sometimes be done through a direct message, or perhaps you can find the person on LinkedIn. If you are lucky, you can get much deeper information in this way, including how they solved their problem and if they are satisfied with that solution.

[9.11] Recipe 5: Fake Button (Smoke Test)

About this experiment

The term ‘Smoke Test’ comes from the early days of computing, when a new piece of equipment was sometimes tested for errors or bugs by simply plugging it in and finding out. If there were sparks and smoke, that was where the problem most likely was.

Today, the word ‘Smoke Test’ is used in lean startup terms to describe an experiment where you simply try to put your prototype on the market, to see if people respond to it.

There are different ways in which you can run a smoke test. For instance, the landing page and the advertising to landing page experiment recipes below are variations. You can do them online and offline, and you can use almost any prototype or product. The limit is your creativity.

What connects these variations is that they all try to get useful data by confronting test subjects with a (fake) product or prototype.

Examples

- Fake Button Experiment (described here)
- Putting it in a store
- Put it on product hunt
- Organize a Demo Event

And even some of the other recipes can be thought of as variations on the Smoke Test:

- Advertisement to Landing Page (Recipe 11)
- Landing page (Recipe 12)
- Crowdfunding (Recipe 15)
- Popup store (Recipe 16)
- Beta launch (Recipe 24)

In this experiment, I’ll use the ‘Fake Button’ experiment as an example, but you can easily come up with your own smoke test.

Fake Button: Add a new button on an existing site for a new feature, and measure how many people click it. The button doesn’t actually work, but leads you to a ‘coming soon’ message, or perhaps to a questionnaire.

Smoke Tests are designed to measure a customer interest or response to a solution or proposition. You’re trying to find out if people ‘get’ it and want to know more.

The Fake Button experiment does this by adding a button on an existing website (with traffic), and measuring if people click on that button. When they do, they get a popup thanking them for their interest (and perhaps prompting them to sign up for a waiting list as an add-on experiment).

If you have a source of traffic, for instance the website of your organization if you're in a corporate venture, you can put a button or link on it and use that to get measurements almost immediately.

Tip: *You can also use emails and newsletters for a fake button experiment.*

Datasheet

Experiment: Smoke Test
Type of experiment: Quantitative
Benefits:

- Easy to do, you don't need a complex prototype
- Good measurements

Dangers:

- You need a source of traffic

Use this experiment to: Find out if people are interested or curious about a proposition, solution, or feature
Time and resources required to set up: Minimal (you'll need to get permission from the guardians of the website though)
Evidence level: Moderate (You have a very simple experiment, but clear measurements. The evidence level is lower because test subjects make no commitment)
Method: Come up with a text for your button, setup the thank you popup, and make sure it's measured. Then launch and watch data come in.
Prototype: A button text and a thank you popup.

Step-by-step

Step 1. Your riskiest assumption

Before you start, you need to figure out what your riskiest assumption is. What do you really want to know? For this type of experiment, you typically can only measure one thing. Think about how you might translate your riskiest assumption into one thing you can measure.

For example, in the early stages of finding Problem-Solution Fit, your riskiest assumption might be along the lines of:

- We can come up with very different solutions that solve the problem for our customers that resonate with our customers

Here, you might think about testing different solutions in terms of a ‘fake button’ experiment, and compare the results.

In a later stage, you’ll be looking for a signal that people really prefer your solution. Here, you can compare either to a benchmark (as explained in section 4.3, Your Hypothesis) or to numbers you researched on a competing product.

Step 2. Define your experiment

Now that you have an idea what you want to measure, and what your riskiest assumption is, create a solid hypothesis. This is a quantitative experiment, so make sure you use the calculator from section 4.3 to come up with the number of visits and clicks you need. If you’re putting your button on an existing page, the webmasters and marketeers in charge should be able to give you a reasonable benchmark for the number of clicks.

Step 3. Dry-run

In many organizations, even placing a simple button on the web page can mean cutting through a lot of red tape. And even if that is not the case, you want to make sure you’re measuring the right thing. Photoshop your new button on the website, and show it to colleagues and (better) prospective customers. What do they think? Did they notice the button? What is their impression? Were they tempted to push it? Why (not)?

Also, once the button is online, test your analytics setup and make sure your button clicks are actually being tracked.

Step 3. Run the experiment

Next, run the experiment and get results. Make sure you download your data once you’re done.

Step 4. Interpret the results

Compare your result with your hypothesis. Can you (in)validate your assumption? Did you get enough data?

When you get a clear result, you know that there at least is a response, and there must be something there. You’ve got smoke, and it cost you hardly anything. To find out more, you can do a more involved experiment.

Hacks

There are many ways to use this test. Besides the aforementioned email list, you can also do it offline. Put up a poster somewhere with a URL or an email address, and see how many

responses you get. Use stickers. Have an airplane fly over with one of those old fashioned banners. Anything to see if people respond. Just make sure you are reaching the right target audience.

***Tip:** it's better if you also know how many people did **not** choose to respond, which is what digital can do for you.*

If you have no traffic source, try to find a partner blog or website that wants to help you out.

[10] Experiments for Problem-Solution Fit

[10.1] What is Problem-Solution Fit

After the initial Idea Validation phase it is time to move on to finding Problem-Solution Fit. This is a crucial stage. And to get a good result, you'll need to let go of your initial ideas about a possible solution, which can be hard.

Just as in the Idea Validation stage, don't go into this trying to prove you're right about your pet solution. It's not about being right. It's not about building the solution that's in your mind. It's about connecting the problem you discovered with a solution that resonates with the people that experience the problem and need it solved. And the solution they resonate with may be very different from what you (from your own limited perspective) had in mind.

> It's not about being right. It's about connecting the problem with a solution that resonates with people.

To make sure you have the best possible shot at coming up with a solution people will love, you need to cast a wide net. You're trying to find the parts of 'solution space' that have the best results. To do that, you'll need to go through a process, poking around in the 'solution space', testing your results with people, and then narrowing the space.

[10.2] Working towards Problem Solution Fit

In the previous stage, you have managed to find people who experience the problem. You have validated your idea resonates with people and found **problem-market fit**.

In this stage, you'll focus on the next step, **problem-solution fit**. Achieving Problem Solution Fit means you'll need to prove that:

1. You are able to find solution options that resonate with customers
2. Customers really prefer your solution over alternatives (Problem Solution Fit)

To do that, you'll first need to create different **potential solutions** for the problem your customers experience, and test those solutions with them to see how they respond.

Then, once you have a better idea on what customers like best in these solutions, you'll need to create a **candidate solution** to test with customers. If you're right, that candidate solution should show clear signals of outperforming alternatives (including the alternatives people may be using currently)

Working towards Problem-Solution fit can be done in many ways, as long as it leads you to a solution your customers prefer. That being said, the following process has proven fast and effective in many situations.

Step 1. Create Potential Solutions

First, it is important to avoid the 'single right solution' pitfall, where you immediately only work on the solution you care for. Doing that would mean introducing a huge bias. You'd be skipping an important step. It's not immediately important to test **your** pet solution, it's first and foremost important to learn **what your customers are looking for** in a solution in the first place. If it turns out that your pet solution checks all the boxes, all the better – but by casting a wider net you'll at least be sure that it does, and that there isn't a better candidate solution you're missing out on.

Computer scientists talk about a 'solution space', an imaginary space where all the possible solutions to a problem live, organized to some principle.

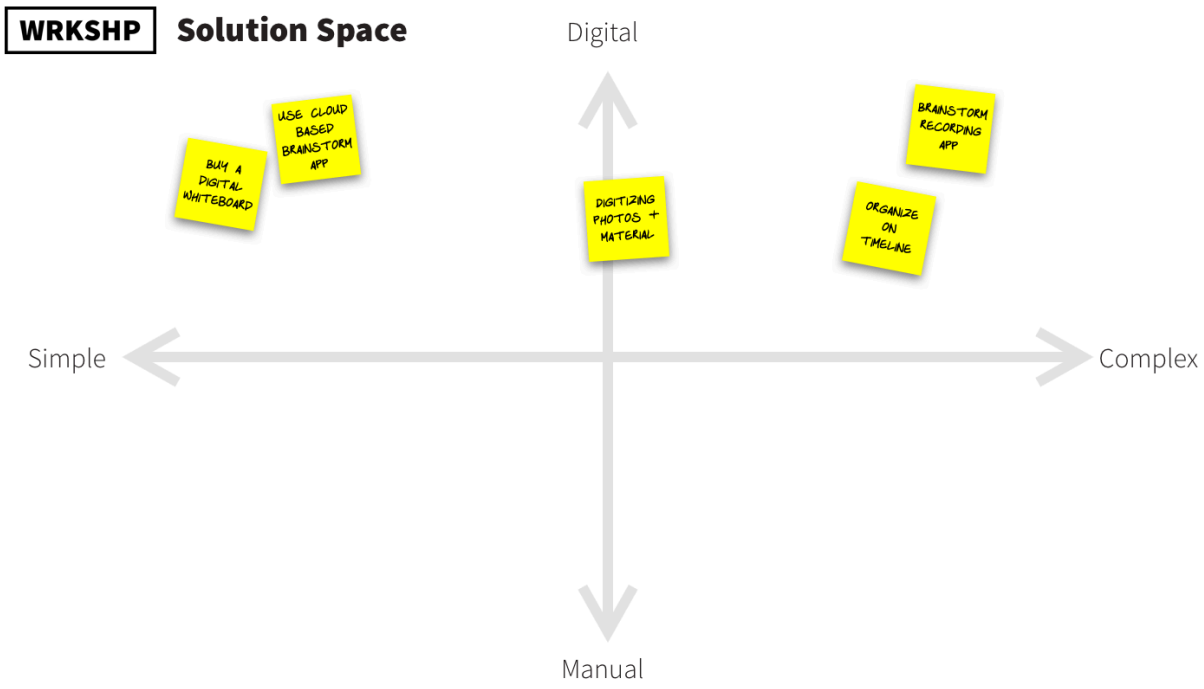
Example
Let's say you're trying to find solutions to the following problem:

"When brainstorming, it is difficult to keep track of all of the material that has been produced. This means we potentially miss out on awesome ideas."

Coming up with some possible solutions leads to the following list of possible solutions:

- An app that records the entire brainstorm
- A way of digitizing photos and material
- An app that is able to organize notes, photos, and other media on a timeline
- Brainstorm in the cloud
- Use a digital whiteboard
- ...

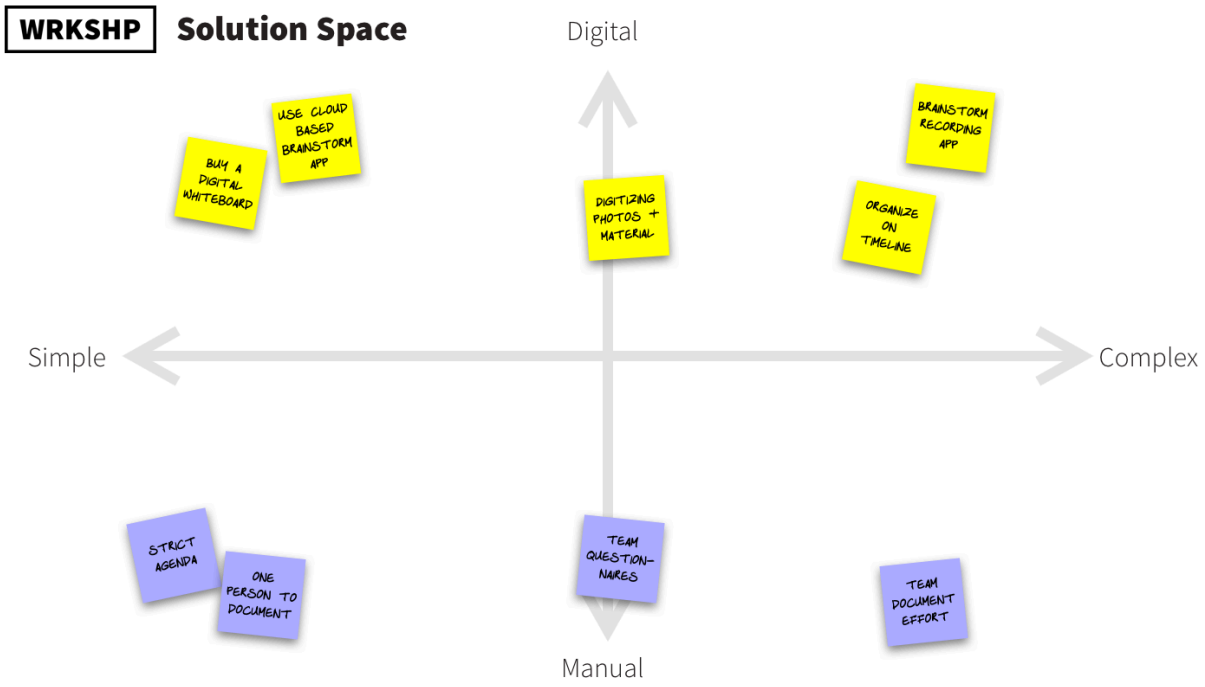
Now, to visualize these in a solution space, imagine a rectangle where on the left you have solutions that are very simple, and on the right solutions that are very complex, on the top you place digital solutions, and on the bottom manual solutions. If you plot the different ideas on this imaginary map, you'll see a pattern. Some areas are more densely populated than others.



In this case, the 'digital' end of the spectrum has been filled, but on the 'manual' side there no ideas have been plotted. That means, you have focused on only a small area of the possible solution space.

If you add more ideas:

- Following a strict procedure or agenda
- Putting one person in charge of documenting everything
- Making all people responsible to take photos and documenting
- Have team members fill out a questionnaire afterward to find improvements
- …



As you can see, these new ideas fill some of the space that was empty before.

Note: *In this example I took only two axes (digital/manual and complex/simple), but in reality there are many more dimensions to explore. You'll need to take that into account when coming up with ideas.*

In this stage, you're not really interested in one solution or the best solution to your problem. You want to explore the solution space, and see how customers respond to solutions from different

To do that, you need to take samples in different areas of the solution space.

Starting from your validated problem, the goal in this first step is to end up with 5–10 'mini propositions' that are all totally different. These will be your 'probes' into different parts of the solution space. Try to come up with a diverse array of ways to solve (parts) of the problem for people.

Think of each mini proposition as a core sample for your solution space.

A fun and quick way to get a ton of ideas for solutions is to use the Wall of Ideas exercise. That will easily leave you with a 100+ ideas to pick from. Of course, 100+ is way too much to test. You'll need to cluster your wall of ideas together and find 5–10 'themes'. For each theme, pick a label.

Example
For a bank, we worked on a new venture around their cybersecurity expertise. They knew they wanted to use that expertise for their small business customers, but had no clear idea how or in what way. We used the mini propositions to come up with directions. Some themes we came up with:

- Trust
- Prevention
- Rescue (after a business has been hit)
- On-demand / 24/7
- Monitoring
- Network
- …

For these themes, we came up with ideas. For 'trust' for instance, we came up with (among other ideas) a certificate (so a business owner can have trust in their level of protection), a way for business owners to inform their own network of threats, and a way to see the security level

of business relations. You can do this for each theme, mentally trying to take care that your themes cover as much of your solution space as possible.

Note: *Unfortunately, the human brain has a lot of problems imagining any space over three dimensions. Because of that it is hard to see if you have missed some dimension in your solution space. The best thing to do is making a list of your themes and going over them to see if you covered enough ground. A trick is to see if you can add the opposite as a theme or idea. For instance, in the example above, 'rescue' is an opposite of 'prevention'.*

Based on these themes, select 5-10 mini propositions. Each mini-proposition is a rough draft of a real value proposition. Make sure each mini proposition is different. For each mini proposition, come up with a unique:

- value proposition (what is it that you sell)
- channel (how it reaches your customer)
- revenue model (how they pay)
- customer relationship (why they will stay with you)
- customer segment or niche (be specific!)
- delivery mechanism (how your product delivers value)
- technology (how it works behind the scenes)

Note: *Don't go overboard, each of these should be one or two lines at most. No business plans!*

A few example mini propositions for the cybersecurity example:

- An 'alarm button' that you can push when you think you've been hacked, where you pay by subscription.
- A certification that you can get after your systems have been tested, which you can use to lower your insurance premiums as a partner deal.
- An 'A-Team' that comes to the rescue to help fix the damage when your systems are compromised, membership plus a number of 'free responses'.
- Education events with inspiring speakers, with ticket sales and paid webinars.
- A dashboard app notifying you of threats, freemium.

It is okay that you can shoot holes in each of these mini propositions and that **they may not be viable business ideas in themselves**. It's okay even if the combination of ingredients that make up the proposition doesn't make total sense.

It's not about *picking* any of these mini propositions as your candidate solution just yet. It's about testing your customers responses to an extreme value proposition, to see what they prefer.

The point is that you can make these 5–10 approaches in a few hours, and then you can go out and test them with real people. And they will tell you which parts they love about these mini propositions, and what they don't understand or dislike.

Flesh out the mini propositions a little, so you can explain what they are about. Make sure to answer the following questions:

- What is it?
- Who is it for?
- Why do you need it?
- When can you use it?
- How will it solve the problem?
- (What does it cost?)

Step 2. Experiment: Test Potential Solutions

Next, run a small experiment. Take your value propositions and go talk to people. It's okay to show them two or three of the mini propositions in succession, and great if they are able to compare them and give their preference. You want to get an idea of what resonates with them.

Present each mini-proposition as an honest solution to their problem, and count the number of people that are enthusiastic about this particular solution. Try to find out what they are enthusiastic about. In this way, you can get information on how they feel about particular **themes**.

Instead of testing and incrementally adjusting on one single proposition **you now have information about a large area of the solution space**, including areas you didn't consider before. Maybe some of those areas are super interesting. And if it is the case that your original 'pet solution' is indeed the one that seems most viable, all the better! At least now you're sure, all at the cost of a few days work. Force yourself to look at the problem from new angles.

Example
In the cybersecurity example above, we (and experts at the bank) had the assumption that a digital dashboard where small business owners could monitor the threat level in their business would be the most successful idea. This dashboard was also something being championed inside the client organization. We included it as one of the solution options, but to our surprise it scored very low. Customers told us they were not waiting for another dashboard with numbers

to keep track of, they wanted to be notified actively when they needed to act. It turned out that a notification system (which we did not really have high hopes for) scored by far the best.

Besides the mini proposition ‘popularity contest’, most of the information you get out of these results will be qualitative and very much open to interpretation. Try to write down only what it is that people say and leave interpretation and drawing conclusions for later. Making sense of the results is best done together with the team.

Experiment Design

To do this step right, you’ll need to setup your hypothesis. Because the mini propositions are easy to test (both face to face and online), you should aim for around 20-30 test subjects at least.

Example Riskiest Assumption:

- You are able to find solution options that resonate with customers

Example Hypothesis: One (or more) aspects of the potential solutions we show to people get a clear positive signal (50%-70% positive)

Example Method: One Pager + interview. For each of the mini propositions, we make a one visual one pager with a fake brand, using photography found on the internet to highlight one key aspect of each mini proposition. We show three of these to our customers, placing them side by side randomly, and ask them to describe what they see. Note the first one they point out. Prompt them to describe what they like or dislike about the propositions.

Step 3. Mashup & Iterate

After getting feedback on the mini propositions with a decent amount of people, it’s time to mix it up.

Look at the parts of the mini propositions that were the most well-received, and try to mash them up into three to five new mini propositions. They could open up new, unexplored parts of the search space that you feel are interesting in the light of the feedback you received, or they could be zooming in on specific aspects that you want to know more about. Again, make them as diverse as possible.

Note that these new mini propositions are again *not the solution* you’ll be looking for in the end! They are another set of ‘probes’ into the solution space that will give you more information.

With the new mini propositions, **do another round of customer interviews**. Try to drill a bit deeper into if and how these solutions might solve the problem they are experiencing.

Step 4. Define a Candidate Solution

Based on the results of steps 2 and 3, piece together your **first draft of a real solution**.

Don't work on it too long, don't overthink it: **it should still be a sketch**. Go over the feedback and try to answer the questions people have. Create a one-page paper prototype or mock-up of what the result could look like. Try to make it look appealing.

Maybe it will take you a few iterations, but in the end **your candidate solution should resonate way better with the target audience than any of the mini propositions**! If that is not the case, then you need to figure out why that is, and go back to fix it.

Step 5. Experiment: Test the Candidate Solution small scale

Create another experiment, and go back to the people you talked to before. Ask them to compare this new solution to the ones you showed before, (and, if applicable, to what they are doing to solve the problem right now, either competing products or workarounds).

Use the Experiment Canvas to map out your experiment, and try to run it with a larger group of people.

Example Risky Assumption: We can come up with a solution people love and prefer over alternatives

Example Hypothesis: More than 70% of 50+ people we interview prefers our solution over alternatives, and wants to signup for a beta test or pilot.

Example Method: Paper prototype or mockup + interview. Ask people to leave their email addresses to stay in touch when they prefer the alternative. Count the email addresses. Ask them to name any friends or colleagues that should also learn about this new product. Alternatively, use a Landing Page and do it online. (See 'Fake it before you make it')

If this hypothesis is validated, you're on to something. And as an extra, you've also started to grow a list of fans that you can build on in the future!

Tip: *When you get a sizable chunk of customers that are asking 'when can I start using it' or 'can we start a pilot project' or 'can I be a launching customer', you're on to something.*

Step 6. Experiment: Prove it on a larger scale

Now that you have evidence for problem solution fit on a small scale, it is time to make sure that the signal holds up on a larger scale, with more potential customers involved.

When this works, you’re ready to move toward the next level.

[10.3] Riskiest Assumptions for Problem Solution Fit

- You are able to find solution options that resonate with customers
- Customers really prefer your solution over alternatives (Problem Solution Fit)

Suitable riskiest assumptions to test in this stage usually revolve around finding out if your candidate solution is really the solution people are looking for.

In the first iterations of working toward problem solution fit, you're learning what customers value most in a potential solution. Here, you're still following a more qualitative approach, and the risky assumptions will likely be very similar.

In the later stages, you're trying to come up with a mix of ingredients for a solution that customers prefer over other solutions. This is where you start to move toward a more quantitative approach. Your riskiest assumptions will become more selective.

Example Riskiest Assumptions (Early Stage)

- We can come up with very different solutions that solve the problem for our customers that resonate with our customers

Or, if you have specific constraints that define the scope of your solution space:

- We can come up with <mobile app> solutions that solve the problem for our customers that resonate with our customers
- We can come up with <scalable> solutions that solve the problem for our customers that resonate with our customers
- We can come up with <ai-based> solutions that solve the problem for our customers that resonate with our customers

Where you can replace what is between <angle brackets> with the type of solution dictated by your design criteria.

Note: *Even though your design criteria may exclude certain solutions, you can still learn a lot from testing excluded solutions in the early stage of your search for Problem Solution Fit. In fact, it is recommended to not limit yourself to only what is dictated by the design criteria at this stage. This will happen anyway when selecting a candidate solution.*

Example Riskiest Assumptions (Later Stage)

- We can come up with a solution people love and prefer over alternatives
- We can come up with a solution people prefer over that of competitor X

- We can come up with a solution that will make people want to switch from their current solution

In the later stage of finding Problem-Solution Fit, you're zooming in on how customers feel about your specific solution candidate. Are they enthusiastic? Do they think this is going to solve their problems? Your riskiest assumptions should reflect this.

[10.4] (Paper) Prototyping

Running experiments and building prototypes go hand in hand. When doing any kind of validation, you need something to show to people in order to get their feedback. That's what prototypes are for.

When validating ideas for startups, products and services, usually people talk of building a prototype, mockup, proof of concept, or even an MVP or Beta as if these are interchangeable. This leads to unclear situations and can mean you lose track of what you're trying to achieve.

If you're not careful, the image that the words 'minimum viable product' conjure up is that of a shiny, usable, 'feature complete' product, ready to be rolled out to actual users. This can put you on a slippery slope: instead of using a lean approach, focusing on validating assumptions, what started as an 'MVP' becomes a goal in itself, and feature creep and gold plating set in.

Taking the name 'minimum viable product' literally , the main differences between a true prototype and an MVP are that a **prototype does 1) *not* need to be functional, and 2) does *not* need to be a viable product**. Understanding this can make your life a lot easier and speed up validation.

> A prototype is nothing but a tool to help you validate a specific assumption.

Validation requires prototypes to be able to create strong and objective signals that tell you if your assumption is validated or not. Asking people about their preferences and interviewing them about what they 'might do' when a product or service became available is simply too unreliable. They can't answer such questions truthfully.

The only way to be at least reasonably sure of what people will actually *do* is by putting them in the situation where the product or service exists, and observing their behaviour.

The function of a prototype is to provide as cheap and simple a way as possible to create that situation.

Therefore, prototypes usually involves faking (aspects of) the actual product. The point is to present users with a situation where they can experience the product or service *as if it already exists,* in order to observe their behaviour. It is *not* the goal to build a 'light' version of your final product.

On the cheap

Prototypes come in different levels of complexity and detail, ranging from ‘lo-fi’ prototypes to ‘hi-fi’ ones. Lo-fi prototypes are far cheaper and faster to make, so it is worthwhile to try to validate assumptions with the lowest fidelity prototypes you can get away with: this means you can iterate faster (and learn faster) and you can try more things.

> It is not necessarily the smallest product imaginable… it is simply the fastest way to get through the Build-Measure-Learn feedback loop with the minimum amount of effort.
>
> — Eric Ries, the Lean Startup

So, rather than looking at the hi-fi side of the range, where MVPs, Betas and technical proof-of-concepts live, it makes sense to have recipes for lo-fi prototypes that you can use to quickly test something.

What prototype should you build?

The answer is: the simplest one you can get away with. Because you’re trying to validate or invalidate an assumption, you don’t always need to prototype your actual solution. If you can come up with a prototype that really helps you to validate an assumption and will give you a clear signal, and that prototype looks nothing like what you finally want to build, that is great.

[10.5] Prototyping Hollywood Style

Prototyping and the Suspension of Disbelief

> “How sophisticated should your prototype be?”

This is a question that you’ll run into whether you run experiments for idea validation, problem-solution fit, or product-market fit.

On the one hand, you need to keep your prototypes as lean and as simple as possible. It doesn’t make sense to spend time and energy on aspects of the prototype that are not necessary to get the feedback you need for the experiment you are running.

On the other hand, it is natural to feel insecure about your product when you’re developing it. This easily translates to insecurity about receiving feedback. What if people don’t like what you are building? For many founders, the natural defense mechanism is to do what they are good at: building. They spend more time polishing, adding features, agonizing about what it should look like — all the time delaying actually showing the prototype to users. This actually increases the risk of building things nobody needs.

Part of the problem seems to come from confusing the prototype with the actual product that you’re building.

That’s right: a prototype and a product are not the same thing.

A prototype does not automatically need the same branding, name, or level of sophistication as the product you’re aiming to build.

The **only purpose** of your prototype is to get feedback to validate or invalidate your current riskiest assumption. Nothing more.

So, you shouldn’t worry about things like branding, design, or functionality, **until** the moment you are actually testing assumptions that focus on these aspects. Before that, it is enough that your prototype looks plausible enough for a person to interact with it. This is what Hollywood calls ’suspension of disbelief’.

> The term **suspension of disbelief** or **willing suspension of disbelief** has been defined as a willingness to suspend one's critical faculties and believe something surreal; sacrifice of realism and logic for the sake of enjoyment.
>
> (Source: Wikipedia)

For a prototype, this means that if the user is able to interact with it in a way that they can take it seriously, and are able to see past the rough edges, it is good enough to test with. It's perfectly all right to create prototypes with a totally different brand and name for an experiment.

Will a prototype that looks closer to the final product give you a better view on what users will actually do? Sure it will. But you don't know yet what that final product should be. So starting out rough and then gradually homing in on a more finished, sophisticated version is the way to go.

Besides, a valuable bonus that a rough prototype can give you is that because it is rough, it leaves things to the imagination. It invites people to volunteer their ideas of what it might become. In their mind, there is room for it to morph into the product or service they are waiting for.

To make your prototype come across as believable, make sure the communication around it is good. When you ask people to take part in a test, make sure that the test is run in a professional way. Hand out clear step by step instructions, tell them what you expect, and what you're aiming for. Start and end any tests on time, and thank them for their input. Doing this will make people much more forgiving when they finally interact with your scrappy prototype.

Start pre-selling the product or service before you have created it to gauge interest. Of course you do need to be careful not to trick buyers into buying something you have no intention of delivering, but if you are upfront about it, this strategy can give a clear idea of the interest for your product or service.

Example: Kickstarter

[10.6] Fake it before you make it

With prototyping for validation experiments, the idea is that instead of first building a functional, complete prototype, you try to ‘fake’ everything as much as possible. Of course, I recommend you to stay within ethical (and legal) limits, and tell people that what they are thinking they are purchasing or signing up for is an experiment.

Faking it can make all the difference, as you’ll be able to learn from your customers before (or while) you are developing your product. This means you’ll have early information allowing you to change direction at a stage where that is still cheap, as opposed to finding out late in the game nobody loves the product you spend millions on.

Examples

Fake Ad: Create a fake advertisement and place it on a screenshot of a website or photoshop it into a magazine. This is a cheap way to give respondents the idea your product or service is legit when you are interviewing them.

Data Sheet: Create a data sheet for your product or service, listing the most important features and characteristics. Show this to potential customers to get a sense of what features they find most important.

Brochure: Create a brochure (digital or using a cheap print-on-demand service) for your product or service. This requires you to come up with plausible copy and branding, but remember that it is only a prototype and that it definitely does not have to be your final branding or name.

Example: A Dutch shared mobility startup needed to find locations where they could place their ‘hubs’ with vehicles. One of the possibilities was to convince real-estate and construction firms to place a hub with their newly developed buildings. The catch: there were no hubs or even vehicles to place there, we just needed to see if there was interest to talk about it. We decided to use brochures of existing housing projects and add in the hubs, as if they were part of the plan. These doctored brochures looked realistic enough to get partners excited.

Physical Mockup: Create a fake version of the product, and show it to people to see their reaction. It can be a digital mockup that you experience on a device, or a physical one. For physical mockups, this allows you to see how a potential customer responds to qualities like size, weight, shape, and colour.

Example: For the book Design A Better Business we created physical mockups to see what colour cover would be most visible in bookstores. We used books with similar size and weight

and stuck on our mockup covers. We placed these in a bookstore and observed. When someone picked them up, we would ask why they did that and what their expectations were.

Paper Prototype: Instead of using programming and design tools to build a digital prototype, simply use paper, scissors, glue, and markers. It allows everyone on the team to contribute, is very cheap to change, and you can show it to people to get their initial reactions quickly.

Example: We used paper prototyping to test whether 'high potentials' were interested in a new financial service. Using just a printer, scissors, and drawing tools, we were able to get a mockup together in hours that was good enough to start a conversation.

Digital Mockup: When wireframes or designs of screens for a digital product or service have been created, convert them into a digital mockup that allows you to observe users interacting with your product. Using online services such as e.g. Maze you can also send these prototypes out to people to gather feedback.

Video Trailer: Create a video trailer of how the product works, and use that to give potential users the idea it is a real product or service.

Example: Dropbox created a video demonstrating their product well before it was finished to capture sign-ups. (Watch founder Drew Houston talk about this experiment on youtube)

[10.7] Recipe 6: Does it Resonate?

About this experiment

Before anything else, you want to know if the problem you're thinking about solving is important to your customers. If they don't have something you can help them with, you won't be able to create a business.

The first thing you can do, almost the smallest experiment imaginable, is to talk to people about your idea and the problem you want to solve. You're trying to find out if they 'get' it. Do they understand where you're going? Can they relate to it? Does the problem resonate?

Datasheet

Experiment: Does it Resonate?
Type of experiment: Validation, quantitative (also qualitative aspects)
Benefits:

- Very easy to do
- Takes little time

Dangers:

- The results you'll get will be 'soft', in that you haven't asked for any commitment.

Use this experiment to: Very quickly find out if your idea has any chance at all, and hear how people feel about it.
Time and resources required to set up: You need test subjects, that's all.
Evidence level: Weak. You are not asking for a commitment and you're not talking to people specifically from the target audience.
Method: Pitch your one minute 'problem statement' to anyone you meet. Ask them if they have ever been in that situation.
Prototype: A problem statement to pitch

Step by Step

Step 1. Define your statement

Write down what you think the problem is that you want to solve. Now, turn it into a simple pitch. Start with something along the lines of: 'have you ever been in the situation where…' or 'have you ever encountered ....'. You can add a personal introduction if you like, e.g. 'I just had this experience where …. Happened. Have you ever had that happen to you?'

Step 2. Run your experiment

Use your pitch with anyone you meet, if at all possible. It can be a bit of a hurdle to get into this habit, but trust me, it will change your life as an entrepreneur. Keep a tally of people that 'got' it, and responded positively that they had the same problem. In those cases, you can go deeper and ask them more about the situation. Write down the key points in their answer afterward.

Also note if people don't 'get' it. If they don't understand, or if they don't know what you're talking about, you may need to change your message. It may be a signal that you are in a niche market, or that people are not ready for your idea. You'll need to find the people who do care about it. If people 'get' it, but don't care about it, you may be looking not at a deep customer need, but at a nice to have. Keep a tally of these situations as well.

Warning: *It may seem so simple that a tally is not necessary. Keep one anyway. Not keeping a tally will make you selectively remember only the answers that you liked, and you'll fall victim to confirmation bias.*

Note: *You're not talking about your solution, only about the problem.*

Step 3. Analyze the data

If you keep this up a number of days or weeks, you'll have some good data to work with. If you have a large number of people that look at you as if you just read to them from the phone book, you know you have work to do. Clearly, you don't speak people's language. If you have a mix of people that are enthusiastic and others that are not, more research is needed. If you can't find anyone that is not enthusiastic, you might be on to something...

The good thing about this experiment is that you can keep it going. It will help you tune your story, and understand how people see the problem, and it doesn't cost you anything at all.

[10.8] Recipe 7: Upvote

About this experiment

Upvote is an experiment where you make clever use of channels to find out if people are interested in a particular solution. Find an online channel, forum, or social media page used by members of your target group. Post one or several questions, or for instance a poll. People will like or upvote answers or vote in your poll. The most popular post wins.

Datasheet

Experiment: Upvote
Type of experiment: Quantitative
Benefits:

- It's easy to do if you have a popular channel
- No investment needed

Dangers:

- If not a lot of people vote the results are low quality
- The people in the channel may not be representative of your target audience

Use this experiment to: Quickly get an insight in how members of the public respond
Time and resources required to set up: Very little
Evidence level: Low, there is no commitment and people answer in a split second..
Method: Create a good post and put it online, see the votes roll in
Prototype: Post

Step-by-step

Step 1. Your assumption

Find your risky assumption and create a hypothesis

Step 2. Write your post

Convert your hypothesis into a good online post, with a nice image and a great headline. You may need to try different versions. Try to create a post where you answer a question, or ask people to like or upvote if they agree. Keep it short and sweet.

If you have different solutions to test, try different posts. Also make 2-3 different versions for each post. In this way, you can try to minimize effects coming from wording, text, and image preferences.

For instance, let's say you want to test mini propositions A, B, and C.

Make three posts for each, A1, A2, A3; B1, B2, B3; C1, C2, C3.

Each has a different image and a different text.

Step 3. Put it online

Select a good channel (or multiple) and put your post online

- Facebook
- Twitter
- Instagram
- LinkedIn Groups
- Reddit
- Quora
- Producthunt
- Specialist forums and pages
- …

Note: *Some of these channels, such as Twitter, Instagram and your own Facebook page, need followers to be effective. If you don't have access to a popular social media channel, try to find a channel that you do have access to or find a friendly influencer to promote your question*

It really helps to spend some time telling people you know about your post, and you may need to post it multiple times on multiple channels to get a response. Try to figure out at what time of day people are online.

Step 4. Interpret the results

For each variation, discard the post with the fewest results. Average the other two posts' likes, upvotes, hearts, or whatever else the channel uses to measure popularity. Compare with your hypothesis.

[10.9] Recipe 8: Test the Competition

About the experiment

Normally, to validate a product, you'd need to take an idea and build a prototype out of it, or some kind of mockup, in order to get feedback.

But what if you don't have a product yet? Or when you only have a half-baked idea? First coming up with a fully working product can be expensive. And although you will definitely get results from that, it is quite different to test with a low-fi prototype or paper mockup than with a real, working product.

So why not test with a real, existing, working product immediately?

Find a product that is in some ways similar (in aesthetic, function, or another respect), buy it, and show it to potential customers. Try not to influence them, but have them explain to you what they think of it, how they use it, what they like, or don't like.

Better still, find existing users of that product that have used it for years. What do they miss? Why would they recommend this product to others? Or why not?

In this way, you can find out how much weight some of the features that only become available once the product is 'finished' will have, and, hopefully, what other products are missing.

Note that you will need to be a bit creative if your idea involves something that does not exist yet. You may need to test several aspects of different products to get a complete picture. The benefit is that you'll be able to start today.

Datasheet

Experiment: Online Research
Type of experiment: Mixed
Benefits:

- Easy to set up
- Relatively fast
- Skip prototyping and building features
- Dive into real user feedback right away

Dangers:

- You’ll only hear about features of an existing product
- People can’t really imagine what it could become

Use this experiment to: Quickly find out what people think about an existing product, validate your assumptions about why they bought or (dis)like that product, find out what they are looking for.
Time and resources required to set up: You need an existing product, possibly existing users of that product, and a location to do 1-on-1 interviews.
Evidence level: Pretty good. You're checking how people actually use an existing product, and what they think about it. You'll have to be careful to carry findings over to your own product, but you'll definitely get a better idea of what people are liking or missing in the current solution.
Method: Display the product, ask them for a response, ask them to use it, interview
Prototype: The competing product.

Step by Step

Step 1. What are you going to explore or measure?

Come up with a list of things you want to learn, and use your riskiest assumptions about how people will use your product to create a hypothesis.

For example, if you assume user friendliness is very important, and you know that great UX design is very expensive, double check that assumption by testing a product people love, which has an (in your opinion) crappy user experience. Is that a problem for the users? Or is the UX a 'nice to have' in this case. What are the aspects of the product people value the most. Is it the price? The feature set? Or something else?

In the end, if you are trying to build a new product that is supposedly a better solution, and you think people should prefer it over existing products, you need to make sure first what they don't like about current products, and second (in another experiment) if your solution makes that deficiency go away.

Step 2. Find a suitable competing product

Look for a working, existing product that has some of the qualities you want to learn about. This could be a competing product, that you want to push out of the market or want to improve upon.

Example: For a mobility startup, we needed to validate if the location of vehicles was an important factor in deciding to register. We located users of three existing car sharing services and interviewed them.

This is an example of experimenting using a competing product. But it could also be a product that has a certain quality you assume is important for your product.

Example: When designing an online trading app for a financial company, we wanted to understand how users might want to see the data presented. Rather than defaulting to the standard graph and spreadsheet type of display, we located three very different ways in which

data was presented: as an animation, a piece of text constructed based on data points, and an alternative type of visualization. We asked test subjects to look at and describe each of these existing apps, to get a better understanding of how important this visualization was for them.

Some ideas for products you can test:

- A competing product you think is very good
- A competing product you assume is not very good
- A product with a similar look and feel
- A product people in your target audience love or use a lot already
- A product that solves the problem in a totally different way

Also prepare a list of key questions to ask to the test subjects, and a list of things to look out for when observing.

Step 3. Run the experiment

When doing an experiment like this, where you observe someone using a product or a piece of equipment, it is very important to not interrupt their thinking process by unnecessary questions.

You want to know what they are thinking, and a good way to do that is to prompt them to think aloud, as they are using the product. Let them explain what they see, why they make certain choices, and how they feel.

Write down your observations and the answers to the test. If possible, capture it on video.

Step 4. Analyze your data

Take your outcomes and compare them with the prediction in your hypothesis. What did you learn?

[10.10] Recipe 9: Customer Journey

About this experiment

Note that you will need to be a bit creative if your idea involves something that does not exist yet. You may need to test several aspects of different products to get a complete picture. The benefit is that you'll be able to start today.

Datasheet

Experiment: Customer Journey
Type of experiment: Explorative, qualitative
Benefits:

- Get an in depth view of the problem from a customer perspective
- Understand where in the life of the customer you can make a difference

Dangers:

- Can be a bit of work to get right

Use this experiment to: Do a deep dive in the world of your target customers. Find out when and how they experience the problem, and what the best moments are to improve their life.
Time and resources required to set up: You need a big sheet of paper and time to do customer interviews
Evidence level: Pretty good, if you're basing your customer journey off of customer interviews.
Method: Have 1-on-1 interviews with customers to map out their personal journey on paper, then combine your findings into one big journey to find patterns.
Prototype: Interview questions

Step by Step

Step 1. What are you going to explore?

Come up with a number of questions about the customer journey of you customers. What are the key moments? When do they experience the problem? What happens before? And after? It can help to map out what you think the journey is, so that you can see how the real examples differ. You want to know what the customer is trying to achieve when they run into the problem. If you are testing risky assumptions, also create a solid hypothesis.

Step 2. Line up interviews

Look for test subjects that have actually been in the situation you're trying to learn about. Schedule 1 on 1 interviews.

Step 3. Run interviews

Together with the subject map out their journey. Ask your test subject what they were trying to achieve in the moment you're interested in. Let them describe the situation. That is your first point on the journey. Next, work your way back in time. Ask them what happened before that moment. How did they end up there? Add new moments on the timeline, and ask the test subject to describe them. What did they do? What did they see? What decisions did they make? How did they feel?

When you've gone back in time to a moment that seems to become less relevant, go back to the key moment and move forward in time. What happened afterwards? Again, add new moments.

When you've got a feeling that you have a good picture, go over it once more with the test subject and check if it's all there.

Step 4. Combine results

When you have a good number of results, try to put everything together. Are the timelines similar? Or are there different possible journeys? Do you see patterns? Is this journey different than you thought? What are the moments where it is the easiest to make an impact?

Keep your combined journey together on the wall or in the cloud using e.g. Miro, so that you can keep refining it and use it in the future.

Hacks: combine and refine with customers

[10.11] Recipe 10: Wizard of Oz

About this experiment

In L. Frank Baum's 'The Wonderful Wizard of Oz', a group of travelers follows a yellow brick road in search of a mighty wizard, the Wizard of Oz. When they finally find him, it turned out that all the wizardry was not real, and that in reality a small man was doing tricks from behind a curtain.

This experiment gets its name from this element of deception. Rather than developing a complex piece of technology first, and testing it on customers later, you'll rig up a prototype so that a human takes the place of that software. The users think they are communicating with a real, sophisticated system, but in reality, unknown to them, a person performs the interactions.

Datasheet

Experiment: Wizard of Oz
Type of experiment: Qualitative
Benefits:

- You get to see the reactions of people to a complex sophisticated piece of technology without having to build it first.

Dangers:

- People might smell you out and see it is a fake

Use this experiment to: Test reactions to a complex and sophisticated system before building it.
Time and resources required to set up: Can take a bit of time, because you'll still need to create a prototype that looks and behaves like the real thing, even if it is controlled by a person
Evidence level: Moderate, people will suspend their disbelief, and you can add a commitment
Method: Build a prototype that looks as if it is a real computer system or piece of equipment, and rig it so it can be controlled by a person in real time
Prototype: The rigged prototype

Step-by-step

Step 1. Your questions

What do you want to learn? Make a list of things in the user interaction you are curious about. What would you like to know? What are your assumptions?

Try to come up with a well defined hypothesis.

Step 2. Build the prototype

Typically, the user-facing side of the prototype will need to be quite sophisticated. Find a designer to help you come up with something that looks plausible and has a decent user interface, if needed.

Spend time to setup scripts, actions, and other things the controller will need to interact with the user in the way a system would. Do they need to see the user? Add a camera and microphone to make it easier.

The test will have to take place in a controlled environment.

Tip: Stream the screen of the user to the people controlling the prototype so they really see what the user is trying to achieve.

Step 3. Test the prototype

Do a number of tests and dry runs before trying it with a real user.

Tip: Create some scenarios for you (and your user) to follow during the experiment

Step 4. Run the experiment

The experiment itself will look very much like a user test. The user interacts with the system, and you observe what they are doing. It's best not to interrupt the experience too much, but it's okay to prompt the test subject to 'think aloud' and tell you what they are seeing, thinking, and trying to achieve.

After the test, prepare some questions to see what the test subject thinks about your system. Was the Wizard of Oz convincing?

Step 5. Interpret the results

Interpret the results, compare to the hypothesis, and validate or invalidate your assumptions.

Example

For a cybersecurity proposition, one of the ideas we proposed was to build a chatbot. We wanted to quickly understand if users would be open to this form of interaction, without building anything. We simply bought a prepaid sim card, setup WhatsApp, and named the account 'Kevin', with a nice robot profile picture. Then we created a script and asked respondents if they would drop Kevin a message. Of course, in reality we were doing the chat ourselves, following the script closely.

[10.12] Recipe 11: Advertisement to Landing Page

About this experiment

A Landing Page (see Recipe 12) is a web page with a single purpose, to convert incoming visitors to (typically) sign up for an email list.

In practice, for some startups that don't have access to an easy traffic source, and use advertisement instead, a landing page can be costly, since you have two conversions to work with.

When you put an advertisement online, it gets shown to people, but only a fraction of those people actually click on the ad. These would then get sent to the landing page, where again only a fraction would sign up. With advertisement, you typically pay for clicks. Every person that arrives at your landing page through advertising will cost you money. Most of them won't sign up, which seems like a waste for your experiment.

It can make more sense to put your experiment on that first conversion rather than the second. That means, you'll make an experiment based on the effectiveness of your ad.

Compared to a landing page this is definitely a lot more tricky. The good thing is that you don't have to pay for all your 'negative samples' (ad impressions), but only for the positive ones (clicks). And you'll need 10 to 20 times fewer clicks, which equates to 10 to 20 times lower experiment cost.

This really starts to add up if you want to test a number of different solutions.

Once you know you get enough clicks, you can always tack on a landing page experiment.

Datasheet

Experiment: Advertisement
Type of experiment: Quantitative
Benefits:

- You find out if people in the target audience are interested in your proposition (they will click on your ad)

Dangers:

- Can be a bit of work to setup
- People don't necessarily need to commit, so it can be hard to see if they clicked by accident or on purpose

Use this experiment to: Gauge if there is interest for different solutions or if one solution is preferred, and as a precursor to a landing page experiment
Time and resources required to set up: Using the right tools you can get it to work in a day or so
Evidence level: Low to moderate. People don't necessarily really commit for something, and larger numbers of visitors and signups give a better signal.
Method: Create online ads (e.g. Google search ads), and see if people click on them
Prototype: Ads

What you need

- Ads. You can create ads in Google Adwords or Facebook Ads
- A benchmark. You need some benchmark to compare your results to, so that you are able to decide if your experiment gives a clear signal.
- A good way to see if your result is significant.

Step 1. Design the experiment

Design your experiment using the experiment canvas, defining your riskiest assumption and falsifiable hypothesis.

Benchmarks

The experiment will try to measure if your landing page is converting significantly better than the benchmark.

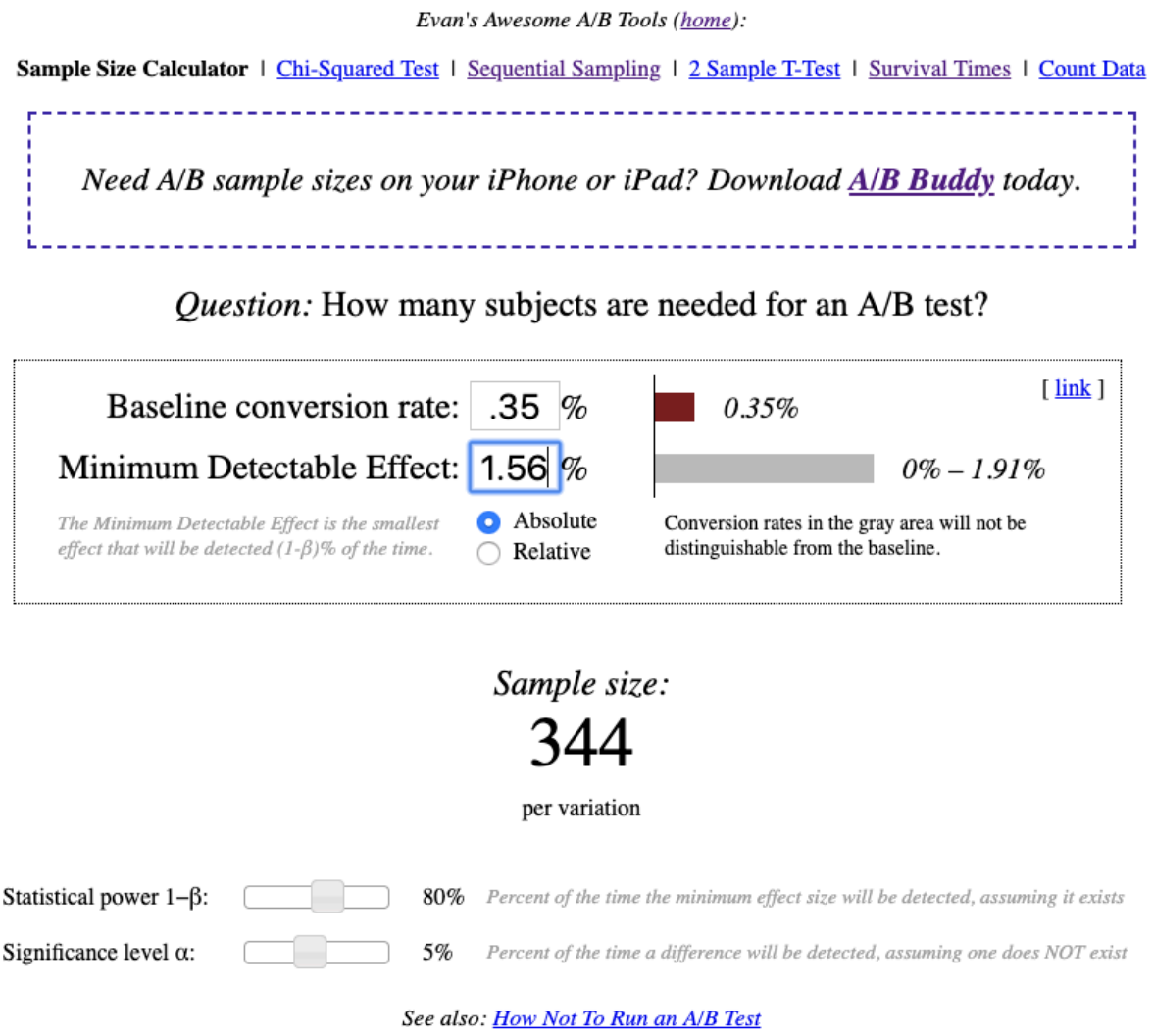
Typical benchmarks for ad conversion:

- Average click through rate (CTR) for a Google Search ad: 0.35% (Source: Google ads)
- Good click through rate (CTR) for a Google Search ad: 1.91% (Source: Google ads)

Note: *Keep in mind that if you are running your experiment with a 'fake brand' or a new brand that nobody knows yet, your rates will likely be lower. Also keep in mind, that a lot of optimization is usually required to get your conversion rate towards the > 1.91% range.*

Traffic

Using the calculator discussed in section 5.4, Your Hypothesis, calculate the number of clicks you need for your ad.



Screenshot of Evan Miller's calculator - See section 5.4, Your Hypothesis, for more explanation

Baseline Conversion Rate

This is where your benchmark goes. I already added in the 0.35% baseline CTR.

Minimum Detectable Effect

The magnitude of the signal you want to be able to pick up. I entered 1.56% here, which means that if your experiment's conversion rate is above 1.91% (0.35+1.56), you will be able to get a statistically significant result. If your experiment's conversion rate lies between 0% and 1.91%, you won't be able to tell the difference from the base rate.

Statistical Power 1-beta

This is usually set to 80%.

Significance level alpha

Typically, it is set to 5%.

What this means

- If you get 344 impressions, and you measure a clickthrough rate > 1.91% (i.e. more than 7 clicks), you can be 80% confident that you have a positive result and you can validate the hypothesis.
- If, after 344 impressions, you do not see a conversion rate > 1.91%, (i.e. less than 7 clicks), you can be 95% confident that you do not have a positive result.

***Note:** If you are testing multiple ads, each ad needs this number of impressions.*

Target Audience

You'll need to know who your target audience is for this experiment, and how to find them online. With ads, targeting is very important. Especially on Facebook there are very precise ways to specify your target audience.

Time

With ads, you can set your daily budget and define your time limit in that way. In this example, you're only looking for 7-10 clicks, which on average would cost you around $25. If you set that as a daily budget, you can get a result in one day.

Note: Due to how Google Ads and Facebook Ads work, the first day can be a bit wonky in terms of measurements. The budget is very low, so I would keep it running a bit longer and get a stronger result.

Your Hypothesis

Now that you have calculated the number of visitors you need and the time needed, you can translate the riskiest assumption in a falsifiable hypothesis.

Example Riskiest Assumption:

- We can come up with very different solutions that solve the problem for our customers in a way that resonates with them.

Falsifiable Hypothesis template:

We believe, that an advertisement experiment
with at least 1376 visitors (4 days with 344 impressions each)
selected from our target audience by running targeted ads,
results in at least 28 visitors signing up (4 days x 7 clicks)
within 4 days

This would cost around $100.

Step 2. Create the ads

Come up with copy

Google Ads and Facebook Ads have excellent examples. Also, look for ads of competitors in Google Search.

Landing Page

You'll need a page to send people to when they click the ad. It's fine to have a single 'thank you' page (perhaps with a signup form). See recipe 12 for tips to create one.

Analytics

Google Ads and Facebook Ads will show you the impressions and clicks for your ad as well as the CTR. You'll have to look the next day, since the (free) analytics don't work in real time.

Step 3. Run the experiment

Launch the ad, and wait for impressions and clicks to roll in.

Keep track of the analytics on a daily basis, and only draw conclusions once you have reached the minimum number of visitors you calculated. Even if you are at a 60% conversion rate, if you haven't reached the number yet it's not actionable data!

Step 4. Interpret the data

Now that the experiment is done, calculate the final CTR and compare it to the benchmark.

Using the numbers from the example above:

- If you get 344 impressions, and you measure a clickthrough rate > 1.91% (i.e. more than 7 clicks), you can be 80% confident that you have a positive result and you can validate the hypothesis.
- If, after 344 impressions, you do not see a conversion rate > 1.91%, (i.e. less than 7 clicks), you can be 95% confident that you do not have a positive result.

If you run multiple days, you can average the result.

Now that you know your result, you're ready for the next step, and you can decide whether to pivot or persevere.

[10.13] Recipe 12: Landing Page

About this experiment

A Landing Page is a web page with a single purpose, to convert incoming visitors to (typically) sign up for an email list. The Landing Page is be setup online, and when a stream of web traffic is directed to the landing page, the conversion rates will tell you if your hypothesis is validated.

There are in essence two types of landing page experiments, one to see if there is any conversion at all (a ‘smoke test’ type experiment) and one that compares the performance of two or more options (an ‘A/B’ type experiment).

The goal of the experiment is to get a clear signal that indicates people favor s specific option you present to them.

Datasheet

Experiment: Landing Page
Type of experiment: Quantitative
Benefits:

- You get potential customers to sign up

Dangers:

- Can be a bit of work to setup
- If you have an unknown product or a new brand initially conversions can be slow

Use this experiment to: Gouge if there is interest for different solutions or if one solution is preferred
Time and resources required to set up: Using the right tools you can get it to work in a day or so
Evidence level: Moderate to high. People really commit for something, and larger numbers of visitors and signups give a better signal
Method: Create an online landing page, send web traffic to the page, and see if people sign up
Prototype: A landing page

What you need

- A landing page. You can use different tools to build an effective landing page (see below)
- A source of traffic. You need to get web traffic to your landing page in order to measure.
- A benchmark. You need some benchmark to compare your results to, so that you are able to decide if your experiment gives a clear signal.
- A good way to see if your result is significant.

Step 1. Design the experiment

Design your experiment using the experiment canvas, defining your riskiest assumption and falsifiable hypothesis.

Benchmarks

The experiment will try to measure if your landing page is converting significantly better than the benchmark.

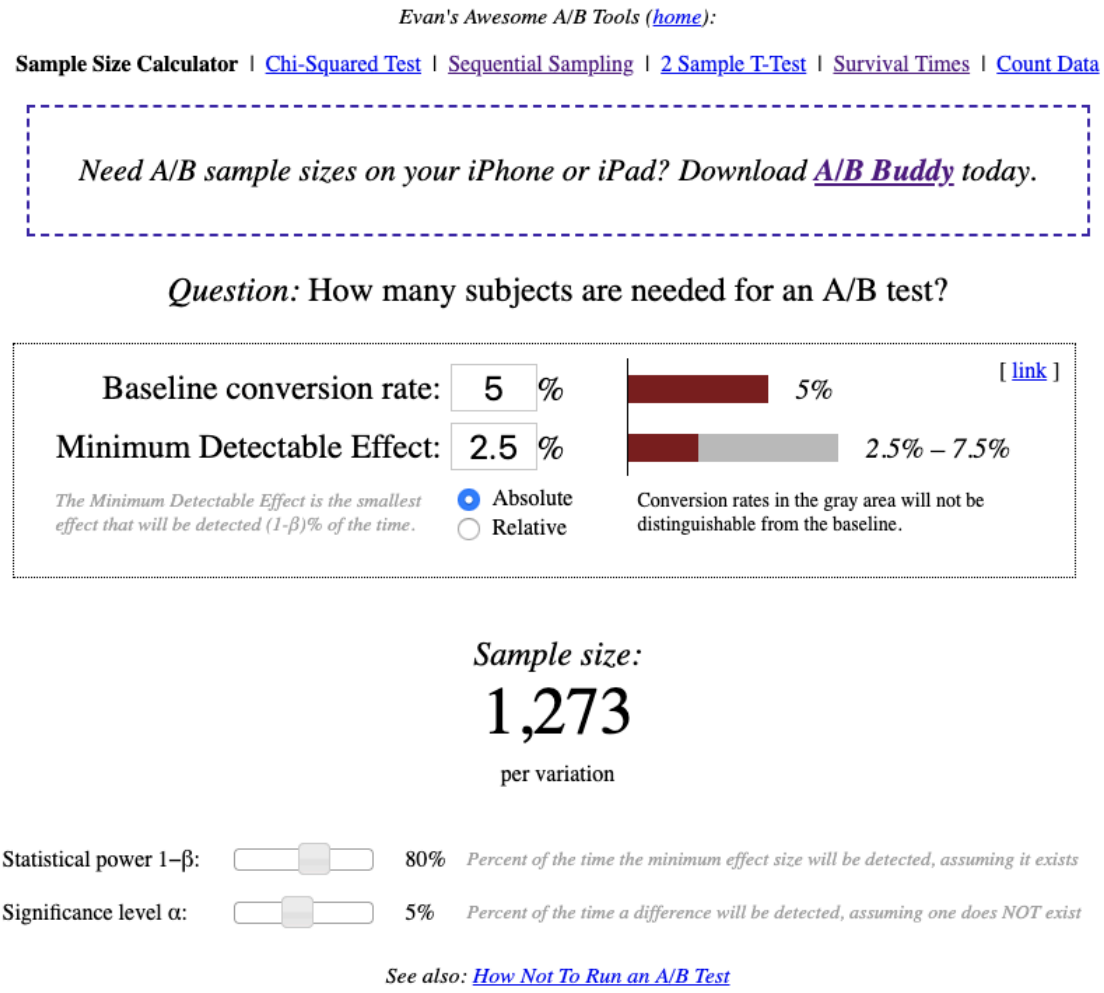
Typical benchmarks for landing page conversion:

- Average conversion rate for a landing page: 2%-5%
- Good conversion rate for a landing page: >10%

Note: *Keep in mind that if you are running your experiment with a 'fake brand' or a new brand that nobody knows yet, your rates will likely be lower. Also keep in mind, that a lot of optimization is usually required to get your conversion rate towards the 10% range.*

Traffic

Using the calculator discussed in section 5.4, Your Hypothesis, calculate the number of visitors you need for your page.



Screenshot of Evan Miller's calculator - See section 5.4, Your Hypothesis, for more explanation

Baseline Conversion Rate

This is where your benchmark goes. I already added in the 5%.

Minimum Detectable Effect

The magnitude of the signal you want to be able to pick up. I entered 2.5% here, which means that if your experiment's conversion rate is below 2.5% (5-2.5) or above 7.5% (5+2.5), you will be able to get a statistically significant result. If your experiment's conversion rate lies between 2.5% and 7.5%, you won't be able to tell the difference from the base rate.

Statistical Power 1-beta

This is usually set to 80%.

Significance level alpha

Typically, it is set to 5%.

What this means

- If you get 1273 visitors, and you measure a conversion rate > 7.5%, you can be 80% confident that you have a positive result and you can validate the hypothesis.
- If, after 1273 visitors, you do not see a conversion rate > 7.5%, you can be 95% confident that you do not have a positive result.

Target Audience

You'll need to know who your target audience is for this experiment, and how to find them online.

Time

You'll need to set a time limit for the experiment. There are two ways to do that:

1. Calculate an expected time to reach the required number of visitors
2. Let it trickle in and see.

I usually prefer the first, but it can be hard to calculate the required time.

Example: traffic from a link on a corporate homepage
Homepage visitors per week: 50,000
Homepage visitors per day: 50,000/7 = +/- 7150
Link clickthrough rate: 1% (get a benchmark from the other links on the page)
Landing page visitors per day: +/- 72
Days needed: 1273 / 72 = +/-18.

A good rule of thumb is to take 1.5 times that number, so in this case roughly a month (27 days).

Your Hypothesis

Now that you have calculated the number of visitors you need and the time needed, you can translate the riskiest assumption in a falsifiable hypothesis.

Example Riskiest Assumption:

- We can come up with very different solutions that solve the problem for our customers in a way that resonate with them.
- This experiment tests one of those different solutions, to see if it solves the problem for our customers in a way that resonates with them.

Falsifiable Hypothesis template:

We believe, that a landing page experiment
with at least 1273 visitors
selected from our target audience by running targeted ads,
results in at least 7.5% visitors signing up
within 14 days

Step 2. Build the Landing Page

Design the Page

A very easy tool to define what goes on the landing page is the Landing Page Canvas. You can find it here.

Page Building

One (expensive) way of building a landing page is to find developers and build it from scratch.

The alternative is to use a landing page tool with a visual editor and put the landing page together in that way. There are lots of landing page builders that can help you out, with nice templates and integrations to speed up the process. Some are free, but most are subscription based. All work just fine for the purpose of a landing page experiment.

Tools

- Squarespace
- Hubspot
- Unbounce

Most have great guides for setting up your landing page in a few hours at most.

An overview of 12 landing page builders: https://blog.useproof.com/best-landing-page-builders/

Email addresses

In many cases, a landing page experiment is geared towards gathering email addresses as a means to gauge the interest for a product or service. So, when you are building your page, you will need a place to store these addresses.

You can use an email service such as Hubspot, Mailchimp, ActiveCampaign, or Drip to capture these addresses, and use the integrations in your page builder to capture signups directly, or use e.g. Zapier to send the email to a google sheet.

Analytics

To calculate the conversion rate, you will need to see how many **unique visitors** your landing page had. The easiest way to track visitors is using google analytics. Your landing page build tool of choice will have a Google Analytics integration, and will most likely come complete with the steps to set it up correctly.

How to Squarespace Google analytics setup

Warning! When you set up analytics, make sure that you **exclude your own visits**. Especially with low numbers of visitors not doing this can really throw your numbers out of whack.

Step 3. Run the experiment

Launch the page, and start sending traffic over using the means stated above.

Keep track of the analytics on a daily basis, and only draw conclusions once you have reached the minimum number of visitors you calculated. Even if you are at a 60% conversion rate, if you haven't reached the number yet it's not actionable data!

Step 4. Interpret the data

Now that the experiment is done, calculate the final conversion rate and compare it to the benchmark.

To get your conversion rate, take the **number of signups** and divide it by the **number of unique visits** from analytics.

Warning! make sure to remove any test signups and such before you do this!)

Using the numbers from the example above:

- If you get 1273 visitors, and you measure a conversion rate > 7.5%, you can be 80% confident that you have a positive result and you can validate the hypothesis.
- If, after 1273 visitors, you do not see a conversion rate > 7.5%, you can be 95% confident that you do not have a positive result.

Now that you know your result, you're ready for the next step, and you can decide whether to pivot or persevere.

[11] Experiments for Product Market Fit

[11.1] What is Product Market Fit?

There are many different conflicting definitions of Product Market Fit. Most are quite ambiguous, and make it difficult to understand when you have achieved it.

> "**Product/market fit** means being in a good **market** with a **product** that can satisfy that **market**."
>
> – Marc Andreessen

While this definition by Marc Andreessen makes perfect sense, it is impossible to measure. What is a 'good market'? What does it mean that the product can 'satisfy' a market?

> If you have to ask whether you have Product/Market Fit, the answer is simple: you don't.
>
> — Eric Ries

The above statement by Eric Ries is profound, but it is no help to you if you are among those that 'had to ask'…

So first of all, we need a definition that allows us to objectively *measure* if Product Market Fit is achieved.

> Achieving product/market fit requires at least 40% of users saying they would be "very disappointed" without your product.
>
> — Sean Ellis (GrowthHackers.com)

40% of users seems like something that is measurable. There are for instance several surveys startups can use to get an idea from their users how they are doing. The question that gets asked is usually a variation of this one:

“How would you feel if you could no longer use product X?”

- *Very disappointed*
- *Slightly disappointed*
- *Couldn’t care less (it’s not that useful)*
- *N.A. (I no longer use product X)*

Asking this to your users will definitely give you an indication of how they think they feel about not being able to use the product. But it is still a question about a hypothetical situation, and as Rob Fitzpatrick teaches us in the Mom Test, people can’t really answer those questions truthfully.

In the end, feedback, surveys, and comments are great to have, but the only real test is behaviour shown while using the product or service.

It would be great to have experiments that you can run besides surveys, right?

How do you know you achieved Product Market Fit?

To test if you have achieved Product Market Fit, you can run several experiments.

The indication that 40% of customers indicate that they would be very disappointed if they no longer have access to a product can be surveyed, but it can also be measured in terms of retention. After all, if people want to continue to use something, they should be seen to continually use it.

This means, that Product Market Fit is for a big part depending on how well you can satisfy your user’s needs, and how happy they are with your solution. This satisfaction is the driver.

AARRR!

To enable measuring retention, it helps to use the AARRR — Pirate Metrics model.The letters AARRR stand for:

- **Acquisition**. You acquire the user. This usually means a sign up.
- **Activation**. The user uses your product. This means a first visit or login.
- **Retention**. The user continues to use your product. This indicates they like your product.
- **Referral**. The user likes your product so much he refers other new users.

- **Revenue**. The user pays you.

For Product Market Fit, especially the Retention and Referral stages of the model are important. It's not per se about revenue yet. It's about how easy it is for you to keep users and to have people refer others to your product or service.

Retention and Referral are the **signals** that indicate the satisfaction level mentioned earlier.

To prepare for your experiments, first map out your AARRR funnel. Try to figure out what you know or don't know about conversion rates. Where do new users come from before they enter the 'acquisition' bucket? How many of them become 'activated'? How many refer the product or service to others?

To prove you have achieved Product Market Fit, you should be able to retain at least 40% of your users as active users. (The period of activity may vary depending on the nature of your product or service, i.e. some things you use every day, others once a month — it's not the frequency that counts, but being the preferred solution).

Steps towards Product Market Fit

Coming from Problem Solution Fit, the first step towards Product Market Fit is to achieve some (very) early traction. Can you go from your ~100 initial ambassadors to ~1000 enthusiastic users? And can you figure out how happy they are to use the product? This is your baseline.

Next, you need to learn how to influence and hopefully increase Retention.Retention is key in terms of Product Market Fit. You need users to be advocates of your product or service, and you must learn how to keep them happy and make them even more enthusiastic.

Finally, you need to learn how to increase Referral rates. Being able to influence this is key to lower your CAC (Cost of Acquiring a Customer) and is a good indication of how happy people are with your service. After all, people hardly tell their friends to use apps they don't really like — even when they might still use those apps if they have no choice.

MVP

Where in the Idea Validation it was enough to have a rough sketch, and you could get away with a prototype for Problem Solution Fit Validation, now you really need to deliver on key features. That means, building a first functional product: an MVP.

There are many confusing definitions of what an MVP is, but I like to think of it as the minimum product that you can build to prove Product Market Fit.

Right after Problem Solution Fit, the first versions of that MVP will be rough (and won't achieve the levels of enthusiasm in your users that you're looking for) — working towards Product Market Fit you'll get a better understanding of what users really look for in the product or service, and your MVP will develop.

Every feature you add or tweak needs to increase user happiness in terms of retention and referral rates. Most new features will start their lives in the form of small experiments, trying to gauge the effect of a change before actually implementing it.

Keep the 'minimal' aspect of MVP in mind! There will be a lot of varied and confusing feedback coming from users, and simply building everything they ask for will not help you. The feedback needs to be deconstructed and you need to look for the biggest common denominators, the real reasons behind the comments. These will show up as risky assumptions when defining your experiments.

[11.2] Designing Experiments for Product Market Fit

How to run a Product Market Fit experiment

Especially in this stage, it is vital to have a good experiment setup and use the experiment canvas as a guide. There are more variables to be aware of than in the earlier stages, and it can be tempting to 'just try' a large number of different things — which can lead to going in circles, not getting clear signals to guide your development, which is the whole point of doing experiments.

The point of running experiments is to get clear signals.

(Note: This doesn't always mean that an experiment should lead to an improvement! It could be that you run an experiment to figure out if a certain feature really is important, and test it by removing it. If the output changes dramatically, now you know that feature is important, and you know it's important to keep supporting it.)

To get clear signals, you need to test the most important hypotheses, and to do that, you need to look at the riskiest assumptions.

It is vital to keep asking *what* users are unhappy about, and *why* they are unhappy about that. Come up with assumptions of the reasons behind what users tell you, behind the behaviours you observe. Use the Riskiest Assumption canvas to sort through your assumptions for the best one to test.

Risky Assumptions

Examples of risky assumptions for product market fit:

- People prefer this solution to competing solutions or workarounds
- People enjoy the solution so much they want their friends to use it
- The ambassadors we found are not 'unique', there are many more potential users we can find
- People come back to use our product and promote it to their go-to solution
- Adding a 'share' button will lead to a higher referral rate
- Adding notifications will lead to higher number of active users, which will lead to higher retention

Hypotheses

Example hypotheses for product market fit:

- 40%+ of the 100+ users we survey indicate that they would be very disappointed if the product or survey was discontinued.
- 40%+ of the 100+ users we survey indicate that they shared or want to share the solution with their friends
- 40%+ of 100+ people we interview are willing to give us 3 email addresses of friends
- 40%+ of our users are active daily/weekly/monthly users
- 40%+ of our users use the 'refer to friend' function in the app
- Adding new feature X increases retention rates by 5% in 1 month

Methods

Example methods for product market fit

- **Face to face interviews** — get a good idea of what they like or hate.
- **User tests** — what are people actually doing? What do they think while using the product?
- **Campaigns** — make campaigns on e.g. facebook that tests how people respond to particular feature sets.
- **Referral experiments**— ask your users to refer others (with or without a reward) and see if they do.
- **Fake features** — Add new features in the product as only a button with a 'coming soon' screen behind it with an option for feedback, and see how many users actually click it.

Keep in mind that the main reason to run experiments in this stage is still not to grow (although it would be nice!) but to learn how you can grow and to increase satisfaction.

The tools to use are a good AARRR model, built in a spreadsheet for instance, that has baselines for retention and referral; the customer journey and personas made earlier (as reference); surveys; user tests; and of course, analytics.

A good analytics setup is vital. You need it to be able to quantify if you are working towards Product Market Fit.

So, if you are working on a digital product or service, it pays to spend time and effort in setting up your analytics across your Pirate Metrics funnel, and to be able to get a grasp of where your users come from, what channels are the best, and what your most happy and loyal users actually do when they use your system. If you don't know that, you are basically flying blind.

Remember that getting from a freshly validated Problem Solution Fit to a solid Product Market Fit will in most cases take quite some time. This is where you find out if your idea and the solution you built on it actually have a shot at becoming the next big thing. If you crack the code of making users fall in love with your product or service you're in business!

[11.13] Recipe 13: Waiting List

About this experiment

The Waiting List experiment is a very good help for startups on their way to product market fit, that want to see if they can get some early traction as well as get feedback about value propositions, feedback, and other things. Starting a waiting list early can help you hit the ground running by the time you want to launch a beta or an MVP.

Some of the most important things for early stage startups are 1) finding good test subjects to do experiments with and learn from, and 2) lining up potential customers to buy the product when it becomes available.

This experiment combines these two goals.

It works similar to a Landing Page experiment (Recipe #12) and an Advertising experiment (Recipe #11), but the follow up is a bit different. Where these two experiments measure the conversion of visitors to (for instance) email signups, the Waiting List continues where they stop. Using your waiting list, you can send follow up information, test engagement through emails, and find out other things about your customers.

Tip: *A Waiting List experiment can start early, even before you have a clear brand (you can always rebrand your landing pages and use the rebranding as another piece of content to excite your fans), and you can keep building it in parallel to all your other experiments and activities. That way, you'll always have a bunch of interested test subjects to get feedback from about your product.*

Datasheet

Experiment: Waiting List
Type of experiment: Quantitative
Benefits:

- Test engagement
- Get potential customers lined up for your product

Dangers:

- You will need to follow up your waiting list campaign and keep people engaged. If you just have people on a list and don't talk to them, they're going to forget about you, or worse: they're going to be frustrated.

Use this experiment to: Find out what features customers like, and build a waiting list of customers.

Time and resources required to set up: You will need to setup a landing page, and mailing list software (see Landing Page experiment)
Evidence level: Moderate, people are not committing to much yet, but they do signal their interest.
Method: Create a waiting list with email addresses, and give out invites or run a simple Facebook campaign. Track how fast you sign up people for the waiting list.
Prototype: (Fake) brand, facebook page, landing page

What you need

- Social Media
    - Facebook Page
    - Twitter
    - Instagram
    - Linkedin
    - Blogs
- Other channels
- Campaigns (if you are in a hurry or have money to spend)
    - (Facebook Ads or Google Ads)
- Landing page
- Mailing list software
- Content for social media and email drip campaigns

Step-by-step

Step 1. Design the experiment

The first step is to define your experiment. In this experiment, the only goal is to get as many interested people on your waiting list as possible. That means, you'll need to learn how to **find** interested people and **convince** them to sign up, and you need to find a way to **qualify** these people as genuinely interested.

What to measure?
During the experiment, you can try different things and see how it affects the **signup rate** and the **quality** of signups. You want to have people on your list that are not just vaguely interested, but ready to buy your product and spread the word.

The average email opt-in rate is 1.95%. Anything above that and you're doing better than half the world.

Evan's Awesome A/B Tools (home):

Sample Size Calculator | Chi-Squared Test | Sequential Sampling | 2 Sample T-Test | Survival Times | Count Data

*Need A/B sample sizes on your iPhone or iPad? Download **A/B Buddy** today.*

Question: How many subjects are needed for an A/B test?

Baseline conversion rate: 1.95 %

Minimum Detectable Effect: 1.95 %

The Minimum Detectable Effect is the smallest effect that will be detected (1-β)% of the time.

Absolute
Relative

[link]

1.95%

0% – 3.9%

Conversion rates in the gray area will not be distinguishable from the baseline.

Sample size:

895

per variation

Statistical power 1−β: 80% *Percent of the time the minimum effect size will be detected, assuming it exists*

Significance level α: 5% *Percent of the time a difference will be detected, assuming one does NOT exist*

See also: How Not To Run an A/B Test

As you'll be in touch through email campaigns, one way to measure this quality is by seeing how often they engage with these emails. What is the open rate for your emails? How often do people click?

Some Industry averages for open rates, clickthrough rates, and unsubscribe rates are:

Industry	Open Rate	Click Rate	Unsubscribe Rate
All non-labeled accounts	21.09%	2.61%	0.23%
Agriculture and Food Services	23.12%	2.69%	0.29%
Architecture and Construction	23.13%	2.55%	0.35%

Arts and Artists	26.03%	2.66%	0.29%
Beauty and Personal Care	17.01%	1.76%	0.31%
Business and Finance	20.47%	2.59%	0.21%
Computers and Electronics	19.39%	1.98%	0.29%
Construction	21.01%	2.03%	0.40%
Consulting	18.96%	2.15%	0.28%
Creative Services/Agency	21.59%	2.51%	0.37%
Daily Deals/E-Coupons	14.92%	2.30%	0.10%
ecommerce	15.66%	2.07%	0.24%
Education and Training	21.80%	2.48%	0.20%
Entertainment and Events	20.41%	2.19%	0.28%
Gambling	18.47%	3.15%	0.14%
Games	19.71%	3.19%	0.24%
Government	26.52%	3.65%	0.13%
Health and Fitness	20.06%	2.18%	0.37%
Hobbies	27.35%	4.78%	0.23%
Home and Garden	22.21%	3.02%	0.36%
Insurance	20.99%	2.09%	0.25%
Legal	21.14%	2.71%	0.22%
Manufacturing	20.51%	2.18%	0.35%
Marketing and Advertising	16.48%	1.74%	0.26%
Media and Publishing	21.92%	4.55%	0.12%
Medical, Dental, and Healthcare	21.09%	2.25%	0.29%
Mobile	18.41%	1.98%	0.34%
Music and Musicians	21.80%	2.68%	0.28%
Non-Profit	24.11%	2.57%	0.20%
Pharmaceuticals	18.95%	2.39%	0.20%

Photo and Video	22.99%	2.90%	0.40%
Politics	22.30%	2.13%	0.21%
Professional Services	20.77%	2.39%	0.30%
Public Relations	20.21%	1.63%	0.19%
Real Estate	19.67%	1.80%	0.29%
Recruitment and Staffing	19.33%	1.81%	0.28%
Religion	25.33%	2.92%	0.13%
Restaurant	20.26%	1.06%	0.28%
Retail	19.36%	2.24%	0.27%
Social Networks and Online Communities	21.13%	3.16%	0.22%
Software and Web App	19.81%	2.05%	0.35%
Sports	23.77%	2.88%	0.28%
Telecommunications	20.27%	2.20%	0.24%
Travel and Transportation	20.03%	2.00%	0.24%

(See the sheet for Recipe #13 for more details)

The average open rate is close to **21%**. The average clickthrough rate is **2.5%**. You can use these numbers as benchmarks and use the calculator from section 5.3 to calculate when you can be sure your list is doing better than average.

For startup waiting lists, I have seen open rates that exceed 50%, and clickthrough rates over 15-20%. That is roughly what you should aim for in the beginning. As your list grows (1000+) the numbers will probably be a bit lower.

When you are running the experiment, make sure you setup your analytics well, so that you can distinguish between different emails, different buttons that are clicked. That way you can run many small experiments at once.

Next, figure out where to find people (see Recipe #11 and #12 for input) and how to reach them.

How many people do you need on your waiting list?
The approach I like to use for this is to work backwards from the number of users you want to start with.

If you want to start with e.g. 100 paying users, and you assume a (benchmark) conversion rate of 5% from your email list (which is filled with potential fans), you'll need 2000 people on your list.

Define your hypothesis, for instance:
"When we launch our MVP (planned 10 weeks from now), we want to have 1500 people on our waiting list, with an average open rate of >30% and an average clickthrough rate of >5%"

Step 2. Plan the experiment

Waiting list experiments can take a bit of time to run, especially if you don't want to use advertisement but want to grow the list organically. Think about what the numbers from your hypothesis mean:

If you want 1500 people to sign up, you probably need to interact with at least 1500/5% = 30,000 people (with a guesstimated signup rate of 5%). You can start with your own network, but you'll probably need to use social media, PR, advertising, or a **smart growth hack** to reach larger numbers.

If you want to only use Google Ads, for instance, this might cost you:

1500 signups / 5% signup rate x $2.69 CPC = $80,700.

That is almost never feasible for startups.

Content marketing, and building an audience on social media first, can help you to get you to 1500 signups without paying $81K to Google. It will mean that you have an extra workload, creating content and interacting with your audience, but it will also immediately start to be a good source of information on what customers need.

Take it one step at a time:

1. Interest: build an audience on social media, blogs, etc (or with ads)
2. Engage: engage the audience (so that they are interested and get to signup)
3. Convert: convert to waiting list (get people to sign up for the list, preferably selectively)
4. Drip: keep waiting list engaged with drip campaigns
5. Run small / additional experiments
    a. Features -- Test engagement on features
    b. Pricing and packages -- Test pricing

Don’t go crazy. If you want to run a 10 week campaign, think about posting twice a week across different channels, and create a content timeline for that. Focus on the core need of your product as a message, but use different angles to build your content. Make sure you post in a consistent rhythm.

Setup a very simple landing page that reinforces that main need (see Recipe #12).

Step 3. Run the experiment

To get your social media channels started, especially if you have no followers yet, try inviting people, following, and posting / reacting on popular channels in the space. Creating messages that trigger people to share or to respond helps.

This is not a social marketing course, so I won’t go in more detail, but there are tons of great resources you can use online to build and grow social media audiences. Below is just one approach you might take for a content strategy.

Divide your content timeline in four week blocks, and these four-week blocks in one week blocks. For each four-week block, select a theme or angle that is close to your product, solution, or the need you are trying to fulfill for your customers.

Note: *You don’t have to come up with everything at the start, you can also add themes as you go along.*

For instance, for a mobility startup, we wanted to build a waiting list while we were developing the product and were running experiments. We looked at various themes along the way:

- The biggest hassles in parking in the city (a previously validated problem)
- Using multiple apps vs just one
- How shared mobility saves the environment
- Shared mobility is cheaper than owning multiple modes of transportation

These themes served as a basis for two short weekly posts, cross-posted on multiple social media. For each post, we created clear UTM codes so that we could track exactly how many people showed up on our landing page (and signed up) from each post.

To get the first people to the page we did run a small targeted Facebook ad campaign, which cost a minimal amount of money, and the founding team shared all posts to their own network. We asked people to tag others that they thought would benefit from the solution.

The question we asked each time was: “Will your neighbourhood be the first to have this new mobility solution? Then signup and tell us why you need it.” - phrased in suitably enticing marketing speak.

In our case, the campaign was very successful, and we received a lot of hits and signups and quickly grew the list towards 1000 people. The proposition clearly triggered people. People also followed the Facebook page and started conversations, which drew in more people.

We also got information about:

- What types of vehicles or modes of transportation would people want to use?
- When do people use a mobility solution?
- What do they do when they use a solution?
- Where do people live that respond?

Because of the targeting in Facebook, it was possible to see from which neighbourhoods and areas in the Netherlands we received the most signups. We used this information to plan beta tests and follow up experiments.

Step 4. Interpret the results

Use the data you gathered to compare with your hypothesis, and see if you validated your assumption that people are interested in your product.

Besides this data, you also have information on what posts and keywords trigger people the most, what the target audience is that engages and clicks, and where people live. Going through this treasure trove of information can give you new insights.

Hacks

- Make it exclusive
- Skip the queue

By making it exclusive, the perceived value of your product goes up. You can do this by giving out invites (making it invite only), or having only a limited number of spots.

A way to learn how interested people are is by adding a commitment, e.g. a payment. In return, they can skip the queue and be the first to try the product when it comes out. You might ask for cash, or for email addresses of colleagues and friends in return. These email addresses are then used to grow your list.

[11.14] Recipe 14: Pre-selling

About this experiment

Pre-selling is a simple experiment that works great for non-digital products as well as digital products. It's a form of 'fake it before you make it' (see 10.6), and it can be used in combination with e.g. Landing Page (Recipe #12) or advertisement, but you can also do it offline by handing out flyers or selling coupons that can be claimed later.

The idea is simple. Rather than first developing a product, you start selling it immediately. Customers can't use the product immediately, but wait for it to come out and pay in advance.

Sometimes, pre-sales are at a discount, or there is a limited number, but if people are really eager to use the product, you may also be able to sell at a premium.

The great thing about pre-sales is that you'll immediately get a strong signal from customers. If they are willing to commit to buying your product in this to them unfavourable situation (they can't use it yet), then you're on to something. The fact that people put in real money makes this a strong signal. It also allows you to experiment with pricing, the value proposition, and other aspects, before you

Warning: *Because you actually do a transaction, this experiment has a few ramifications to be aware of. Make sure people understand it is a pre-sale. Also, make sure you keep track of sales, so you know how much product you need to deliver and where to ship it. Check your local laws for this situation, and what you need to do to protect your customers and yourself. What will happen if you can't deliver the product? What will happen if the final product turns out to be more expensive? What if it is not delivered on time? Make sure you're aware of this, as it has happened that startups folded because of a failed pre-sales attempt or got into the news because of it. Remember the story of Fyre Festival? So be ethical, and be prepared to pay back the money if you change your plans!*

Datasheet

Experiment: Pre-selling
Type of experiment: Quantitative
Benefits:

- You get information on some of the most important aspects of your product, namely pricing and demand, before you build it.
- Based on this information you're able to adjust your direction for what you're building.
- You get cash flow.

Dangers:

- You're entering in a transaction, which means you have to uphold your end of the deal. If you don't, there can be reputation damage or worse.
- As soon as you have pre-sold, you're committed to deliver. It becomes harder to completely change direction or postpone delivery with a horde of angry customers at your front door.
- Spending the money before you know you can deliver is dangerous

Use this experiment to: Validate demand and pricing.
Time and resources required to set up: Limited, you need a simple prototype or specification to show before you can sell it (see 10.6 Fake it before you make it), and a channel to sell it through
Evidence level: High. People paying you even when they can't use the product immediately or have tried it means they really want to have it.
Method: A simple prototype, datasheet, or landing page for the product, a price, and a sales method.
Prototype: Anything from 10.6 Fake it before you make it.

What you need

- A way to keep track of who bought something so they can claim the product when it is out and you can notify them of any updates.
- An idea about the price
- A simple prototype
- A channel to sell through

Step-by-step

Step 1. Define your experiment
The first step is to define what you want to validate. This experiment will tell you if the given pricing and feature set of your product are enough to make people buy it, even if it is not ready. But how many pre-sales do you need?

There are two ways to think about this.

1. You could calculate how much it will cost you to build the product and try to get that money together by pre-selling.
2. You could calculate a number of pre-sales that would signal you have an above average or very good product.

In both cases, the speed and quantity of pre-sales will tell you something about your product.

Benchmarks

It is a bit harder to give benchmarks for this, because it can really depend on your product and your market. A digital download will be very different from a high-end piece of hardware. It pays off to get some information from your market.

Look at other (similar, or competing) products in your market. Ask yourself:

1. How much do they cost?
2. How well do they sell?
3. Where can people buy them?

It may take you a bit of detective work to get an idea of how well they sell as people tend to keep this to themselves.

For some types of products and sales channels this information is available. If you're looking to build mobile apps, for instance, or if you use platforms such as spotify, they may be able to give you some insights in sales numbers or conversion rates.

Can't find a decent benchmark? Then approach it from the other side. Look at it from a conversion rate point of view.

To make this experiment work, you'll need to bring your offering to people's attention. You might use ads for this, but it is way better to build your own audience (see Recipe #13: Waiting List). Make sure the people you target are selected for their interest in this specific product. Take the number of people, and use a conversion rate benchmark of > 5% of that group of people.

If you have an existing mailing list, for instance, you probably won't reach 5% conversion on that entire list. Instead, start with a campaign to filter the people that are really interested in the problem you're solving and in your solution. Use that subset for your experiment.

Using these numbers construct your hypothesis.

Step 2. Prepare your experiment
Create the materials you'll need to sell your product.

- A 'fake' prototype (see 10.6, fake it before you make it).
- Copy and visuals for your product
- A datasheet with the features
- A price
- A sales page
- Legal and finance
    - Disclaimer and notice that this is a pre-sale
    - Check how you should deal with tax and VAT
    - Setup a way to pay: for example use Stripe

    - Receipts
- Records
    - A way to record your sales and relevant data so people can claim their product when it's done.
- Setup email automation to keep your customers informed, e.g. through Drip or Mailchimp.
- A list of channels to plug your product

For digital products, think of places like Producthunt or Betapage to launch your presale. You can also use other platforms such as Etsy, Ebay, Craigslist, and of course Facebook and other social media channels. Try to get endorsements from friendly influencers. Have a look at the Pirate Metrics canvas in Tools and think of a way to get people into the funnel.

For offline products, you can be creative. Think of putting your product (with permission) on a spot in a friendly store, have a demo event, go to meetups, and put up posters and stickers.

Examples

Event tickets
For a startup that wanted to launch a new type of networking event, we pre-sold three different types of event tickets at several meetups and events where they assumed their target audience got together. We did not manage to sell large numbers, around 20 or so, but the conversations about the product and the strong preference for one of the three tickets (based on the price and duration of the event) was a good signal to go ahead and start organizing the first event.

Tesla
Tesla pre-sold their first car, the roadster, starting in 2006. They had only a prototype ready at the time. The first 100 roadsters sold out in three weeks, but would only be delivered in 2009. Tesla have been using this strategy for all of their new models.

[11.15] Recipe 15: Crowdfunding

About this experiment

Crowdfunding is a form of pre-selling (see Recipe #14), where (usually) a crowdfunding platform (such as e.g. Kickstarter) is used to present a product to interested backers before it has finished development.

When crowdfunding first became popular, the platforms had to deal with a number of untrustworthy and failed crowdfunding projects, and as a result most platforms now have very strict regulations on how to run a campaign.
Crowdfunding can still be beneficial, but the additional load of adhering to these regulations has made it less interesting. The commitment from the side of the startup can be great.

It is also possible to run a crowdfunding experiment without using a platform. In such a case, you can get people from your own network as backers for your product. For instance, you could see if you can get 20 people from your network to back you with 50€ each, to see how they respond to your proposition.

Compared to an actual pre-sale, crowdfunding campaigns make less powerful experiments. The reason is that people may back your product that do not intend to use it themselves, but want to see it realized. The predictive power of a successful crowdfunding campaign may therefore be less predictive of how actual customers may behave.

In most situations where you could run a crowdfunding campaign, you could also run a pre-sales campaign, which is easier and gives better evidence. Crowdfunding campaigns are more effective to get funding than as an experiment.

Tip: *Before you attempt this, talk to founders who have (successfully or unsuccessfully) ran a crowdfunding campaign. What did they learn? What do they think of your idea to run this experiment?*

Datasheet

Experiment: Crowdfunding
Type of experiment: Quantitative
Benefits:

- You can use crowdfunding platforms as channels to reach potential backers if you don't have channels to reach customers.
- Platforms can protect you against problems due to failed crowdfunding campaigns, because they hold the backers' money in escrow until the kickstarter succeeds.

Dangers:

- It can be a lot of work to comply with platform rules.
- Launching a campaign definitely does not guarantee visibility, you'll need to do additional marketing.

Use this experiment to: See if early adopters and crowdfunding backers believe in your product.
Time and resources required to set up: Extensive. Crowdfunding platforms have strict rules for documentation, video, and community management
Evidence level: Moderate to high. The backers may not be representative of your target audience.
Method: Create and run a crowdfunding campaign, create marketing materials, communicate with backers
Prototype: Campaign and marketing materials

Step-by-step

Step 1. Define your experiment
The first step is to define your experiment. Find relevant benchmarks by looking at failed and successful crowdfunding campaigns for similar or competing products aimed at the same audience. Try to figure out what their success rate is, and how much money they aimed at. How fast did they achieve their goal?

To use your crowdfunding campaign as an experiment, you'll need to be able to compare its success with other campaigns. You can compare the rate of new backers coming in, the time needed to reach the goal, and the amount of money raised.

Use the numbers you have found to come up with an appropriate hypothesis.

Step 2. Prepare your experiment
Prepare the materials you need to run your campaign. Make sure you have talked to people with experience running crowdfunding campaigns.

Create your project's page on the platform and fill it with content. Look at other (successful) campaigns to get an idea of the content you need.

Setup and start your marketing. Notify people in your network of your campaign.

Step 3. Run your experiment
Run the experiment, and keep your backers up to date. Keep adding new content and doing marketing.

Step 4. Interpret the results

Once the campaign is over, it will either have reached the goal, or it won’t. This can be your signal.

[11.16] Recipe 16: Popup Store

About this experiment

Another way to validate if people will buy your product is through a popup store. A popup store only has to be in operation for a limited time, and it is much easier to run a shop for a week than it is to operate year round. Pop-up shops are a great way to see if there is interest for something before going all in.

Pop-up stores can be run inside a friendly established store (as a shop-in-shop), at a friendly venue that attracts your target audience, or in a temporary space in a good shopping location, such as a building scheduled to be demolished or refurbished or an unused space. It's also possible to create a mobile pop-up store and use e.g. a truck or stall.

Because pop-up stores are temporary, they can be a great way to try a new product with a new audience. You can even move the store and test if it works in different locations.

Setting up a popup store can be a bit of work. You need a space, actual products to sell, a cash register, you'll need to decorate (even if it is minimal), and you'll need staff. In most countries, you'll also need to get a permit and comply with other regulations. This means it is usually not feasible to run a pop-up shop for just a few days. A typical duration is a few weeks to a few months.

A good way to overcome some of the problems related to operating a brick and mortar shop, especially if you have no experience doing so, is to partner up with a store or partners that do have this experience. Find an experience retailer that can help you run the shop on a daily basis.

Datasheet

Experiment: Pop-up store
Type of experiment: Quantitative (but you'll get qualitative feedback as well)
Benefits:

- It is a great test to see how people respond to your product
- You're able to observe and talk to customers
- You can test if the product is in demand

Dangers:

- It is quite a bit of work to set up
- Even on a shoestring budget it will still cost time and money
- You need an actual product to sell

Use this experiment to: Get feedback about pricing and features from people in a real-world store
Time and resources required to set up: Involved
Evidence level: High. People either like and buy the product, or they don't
Method: Open a pop-up store and start selling product
Prototype: The store and the first sold products

What you'll need

- A location
- Products to sell
- The necessary permits
- A cash register
- Furniture
- Staff
- Marketing materials

Step-by-step

Step 1. Define the experiment
The first step is to define what you want to validate and learn about. It can be difficult to get a benchmark for this, and it can help to talk with relations or partners that have retail experience to see what would be a good number of sales.

In a good store, the **average conversion rate** ranges between **20%** and **40%** for most retailers. Using that average, that means about **70%** of shoppers are leaving the store without buying anything. These numbers are based on established stores with many products. Your pop-up store may be new and interesting, and attract people, but you won't have the benefit of an established brand or product. Aim for a rate in this region.

Another metric you can use is how many visitors enter your store. You will need to track this to calculate a conversion rate, but it can be a target by itself. To get a benchmark, spend some time observing people entering neighboring stores (or ask them).

Step 2. Prepare and run
This means setting up the store, getting permits, and decorating. Advertise an opening date for your pop-up shop. Get everything ready, and organize an opening party.

Invest in technology to measure the number of people entering the store. A cheap way is to keep tallies at intervals, and extrapolating. You can also use electronic counters.

When you run the experiment, make a habit of talking to customers, finding out what made them come into the store, and why they have or haven't made a purchase.

Step 3. Interpret the result
At regular interviews and at the end of the experiment, calculate your conversion rates. Based on all the feedback you have gathered, decide how to proceed.

Hacks

A great hack is to sell similar or competing products in the store, side to side with your own product, or even instead of your own product. This will give you a lot of information about how your product compares. (See also Recipe #8: Test the competition)

ADD: https://www.storedna.co/assets/pdf/Future_Of_Store_Innovation_Samsung.pdf

[11.17] Recipe 17: Concierge Model

About this experiment

In a Concierge Model experiment, a human manually fulfils a (large) part of the interaction.

In contrast to a Wizard of Oz experiment (see Recipe #10) the potential customer doesn't need to think they are talking to a computer, however.

By using people to carry out the behind-the-scenes processes, taking the place of complex logistics systems, you'll learn about customer needs, expectations, and how they interact with your product before worrying about scaling and optimizing for efficiency.

Removing the need for complex systems, doing the work yourself, will make you familiar with the details of the process, and reduce the time and cost needed to set up.

> Startups should do things that don't scale
>
> -- Paul Graham

Datasheet

Experiment: Concierge experiment
Type of experiment: Mixed
Benefits:

- Learn about everything around your product and the way customers purchase and interact with it before investing in a scalable process

Dangers:

- It's not scalable so you'll need to be able to transition to a scalable model at some point. If the experiment is successful, its success can become a distraction from that transition.

Use this experiment to: Learn about everything around your product and the way customers purchase and interact with it before investing in a scalable process.
Time and resources required to set up: Quite minimal.
Evidence level: Good. People make the commitment to buy your product.
Method: Sell a product and do all logistics and handling to deliver and support the problem by hand.
Prototype: A product to sell

Step-by-step

Step 1. Define the experiment

First, define an experiment. Besides actual selling, you can measure things such as the amount of customer support required and aspects of the delivery process. Come up with a hypothesis to match this.

Step 2. Prepare the experiment
Set up the experiment, and find channels to sell your product. See experiment #11 (advertising), and #12 (Landing page), for inspiration.

Figure out how you are going to handle incoming orders.

For example, Zappos founder Tony Hsieh bought the shoes he offered online in local retailers. Dutch startup 3D Hubs initially served their customers by finding owners of 3D printers in Amsterdam willing to take paid printing jobs.

Step 3. Run the experiment
Run the experiment, and handle orders as they come in. Try to find opportunities to talk to customers about their experience.

Step 4. Interpret results
Interpret the results and compare them with your hypothesis.

[11.18] Recipe 18: Free Drives Paid

About this experiment

The idea of this experiment recipe is that you get customers to use (a basic version of) your product for free, and then experiment with that customer group to see what features are key for them to make the step to paying for your service.

It is based on a strategy Spotify applied to get non-paying customers to become paying customers. Spotify initially had the assumption that people would transfer from free to paid automatically in certain percentages after experiencing the service. After launch, they found out that the percentage of users switching to a paid subscription was very low. This was a problem. Adding to this, they found out that they were not able to make extra income through advertising.

This meant they were stuck with a huge group of users that listened to music for free, without ever transitioning into paying customers.

In their experiment, they tried various things to change this. One example was that they used annoying ads to make people transfer, when they found out people really hated their music experience to be interrupted with ads. Although the ads still didn't bring in extra income, they could be used to drive people to the ad-free paid service. Another feature they tested with to drive paid subscriptions was mobile service. Downloading songs on your mobile for offline use, which was (or still is) vital in many less-connected areas in the world at the time, was only possible with a paid subscription. These experiments helped people become used to paying for their Spotify account.

A free-drives-paid experiment has the added benefit of getting users to sign up and use (a basic version of) your product for free, testing the Pirate Metrics steps of acquisition and activation. These users will give you a lot of information on how they use the product and what their behaviour is, and give you a user base that you can use for further tests, just like the Waiting List recipe does.

This type of marketing or community building through a free tool is often called 'engineering as marketing'. You have undoubtedly encountered examples online.

Example
The Startup Calculator I helped develop was set up as a basic version of a more complex tool, helping startups to explore how profitable they would be in different growth scenarios. We launched the free tool in order to build a large list of potential customers we could use to test with. We sent them surveys and interviewed them on things we wanted to learn before committing to a paid tool.

The fact that you need a working (free) version means the experiment itself can only start once you have built that version, but it can still be very small or rudimentary. Use the ‘fake it before you make it’ strategy to come up with cheap solutions.

Datasheet

Experiment: Free Drives Paid
Type of experiment: Quantitative
Benefits:

- Build a user base
- Don’t immediately run into a pricing wall when people don’t want to pay for the product yet because it is too unfinished
- Measure which features and pricing are most important drivers for people to pay for the product

Dangers:

- You need a working free version which can take some time and attention

Use this experiment to: Measure which features and pricing are most important drivers for people to pay for the product
Time and resources required to set up: Building the free version can take some time
Evidence level: High. Once you have a large group of free users and can get people to convert to the paid version you are measuring real commitments.
Method: Launch a free basic version of the product, get a user base, and try to convert them to a paid service
Prototype: Free basic version with an upsell to a more complex version

Step-by-step

Step 1. Get free users
The experiment has two steps. One is launching the basic version and getting a number of free users in. This can already be a first hypothesis to test. Look at recipes #12 (Landing Page) and #13 (Waiting List) for inspiration.

Step 2. Get feedback
The second phase starts once you have a significant number of free users. Use the sample size calculator to find out how many (it will probably be around 250-500). Use interviews and surveys to get information on what features free users are really missing, and measure what features they use the most. Come up with a potential new feature for a paid subscription, and consider limiting the feature people use most in the free version. Test your assumption first qualitatively in conversation with a few of your users. If the results are positive, build the new feature.

Step 3. Get paid users

Define your experiment. How many people will convert? Run the experiment and launch the new subscription model. After running the experiment, interpret the results.

[11.19] Recipe 19: Viral Effect

About this experiment

If you manage to get your product go viral, that is a sure measure that it may become popular in your customer segments. It also proves that you are able to acquire new customers and that you may be able to get people to refer your product to others.

However, it is hard to 'make' something go crazy viral. There are not a lot of recipes. Therefore, this recipe is not so much about creating the next meme hype as it is about getting people to share your posts, ads, and links about your product. The goal is to create items that are shared on social media and online. If you manage to do that, you'll be able to get in more interested potential customers.

This experiment combines well with the Waiting List (Recipe #13), Landing Page (#12) and Pre-Selling (#14). In a way, it is similar to Upvote (#7) but more involved.

Datasheet

Experiment: Viral Effect
Type of experiment: Quantitative
Benefits:

- You'll learn how to get more interested customers to your product, for 'free'.
- You'll learn about what interests people in your customer segment
- You'll learn how to make people become referrers.

Dangers:

- It's not easy to make something go viral with a click of a button.
- Access to high-traffic channels is a factor
- It can be tempting to focus on the viral factor and lose sight of marketing your product

Use this experiment to: Learn about what interests people in your customer segment so much that they want to share it
Time and resources required to set up: Can be extensive, you'll need to create multiple pieces of content, do several tests, and 'seed' the content on multiple channels
Evidence level: Moderate, people may share your content for different reasons than that they are interested in the product
Method: Create and post content, measure if it gets shared a lot
Prototype: Your 'viral' posts and content
Side effects: Extra potential customers

Step-by-step

Step 1. Define your experiment
Define your experiment. You can measure two things:

1) The number of likes, shares, retweets, and other engagements on the original posts
2) Click throughs to your website

If people start to share your content, you may lose track of all the places they repost it. So measuring only the original post is of limited value. Combining it with click throughs can help to get a better handle on this.

Find out what benchmarks are.

Examples:
LinkedIn
I came across this interesting analysis of a post going viral on LinkedIn. It has interesting hypotheses about what made it go viral, as well as data on the amount of shares and likes

Medium
An analysis of a post going viral on Medium
How many claps do great Medium stories get?

Other:
An analysis of 4 posts that went viral - examples 3 and 4 give a clear idea of what could happen to your traffic if it works.

Step 2. Create and test content
Focus on high quality, relatable content.

Start by looking at other posts in your customer segment that have high numbers of shares. What patterns can your see? How do people respond?

Next, brainstorm at least 5-10 different directions for your post. Go from whimsical to to the point, and from abstract to practical. Use the patterns you uncovered. Test your ideas with some selected members from your customer segment or waiting list. Is there a direction they prefer? Select the top 3-5 posts.

Then, find out how to 'hack' the channels you use. Understanding the algorithm the channel uses to select posts to show to users is vital. There are many tutorials and how-to's online on how to do this for e.g. LinkedIn and Facebook. Use the lessons from these tutorials when finalizing your content.

Finally, finish your posts using all the pointers you have. Test the final results again with people from the audience.

Step 3. Seed content
Start what you hope will become a wildfire by ‘seeding’ your content on strategic channels. Try to find friendly blogs and influencers to repost or link your content. Get friends and colleagues to like the posts when they go live.

Many algorithms are time-sensitive, and promote new content. Find out what the best times are to get your content online and coordinate your seed strategies with that.

Step 4. Babysit
Once the initial round of likes and shares is over, it’s your job to keep fueling the fire. Repost, promote, and otherwise bring your content back to the viewer’s attention. If you had initial success, use that to convince blogs and media to cover it.

Step 5. Interpret results
Follow closely what is going on. After a week or so, you can start to add up the results. Did you see a continued raised influx of people to your site? Or was it a spike and then a sharp drop off?

Also, search for the comments, reposts, and other things people did with your content. What did they say? Was the context positive? Or negative?

Getting a post to go viral is hard, and you’ll most likely have to try it multiple times. Every time you’ll be learning about how you can market and sell your product.

[11.20] Recipe 20: Network Effect

About this experiment

The Network Effect experiment is similar to the Viral Effect (Recipe #19), but here, you're looking at word of mouth, offline, and in network shares.

The idea is to get people to refer your product to their friends and colleagues.

Example
A great example of this is Dropbox. They used one version of a network effect experiment when they launched: they gave away free storage if your friends would join Dropbox through a friend link they supplied. Another way in which they use a network effect is that you can share Dropbox folders, but to use them you need to become a member.

You can come up with an experiment for your product where you measure a network effect, for instance:

1) Give people invite codes to share
2) Ask people to give you three email addresses of friends / colleagues that would benefit from your product
3) Give out special friend codes to your users, that friends can enter when they sign up; in this way you can track how they signed up
4) Ask them if they were referred when they sign up
5) Hand new members physical coupons to hand out to friends
6) …

In all cases, the point is that you need to be able to track how many people sign up because others referred them.

Warning: *If you reward people for a referral, you will inflate your referral numbers. This means you'll get a lot of bogus referrals or people that are not in your target audience. When you're designing your experiment, try to minimize this risk, and ask people to refer only those friends and colleagues that would benefit the most from your service. You're running an experiment, not doing marketing.*

Datasheet

Experiment: Network Effect
Type of experiment: Quantitative
Benefits:

- Learn if your product is good enough that people want to share it with their friends and network

- Get more interested potential customers

Dangers:

- If you reward people for referring, you might inflate your score

Use this experiment to: Learn if your product is good enough that people want to share it with their friends and network
Time and resources required to set up: Can be minimal
Evidence level: High, if the experiment is set up right: people usually only involve their network if they have no risk of reputation damage, i.e. if they think your product is really good. If they don't think that, they won't take that risk.
Method: Get people to refer your product to other people they know
Prototype: A functionality that allows people to refer easily

Step-by-step

Step 1. Define your experiment
Design your experiment. Try to come up with a benchmark for referrals.

Some benchmarks that can be found for existing referral programs (where customers can refer purchases to other customers, resulting in a sale) can be found in this article.

Some highlights:

- 36% of customers were actively referring, 64% were not actively referring
- 61% of active customers made 1 referral (active customers are customers that made at least 1 referral)
- 34% of active customers made 2-10 referrals
- 5% of active customers made 11+ referrals
- On average over a year period a customer produced 0.87 referral leads.

What does this mean? Well, first, these numbers were for people enrolled in a referral program, i.e. they were self-selected for wanting to refer. That means, your numbers may be lower. Of those, 36% of customers referred at least 1 other customer.

That is something you could use as a benchmark for a hypothesis. Can your (new, unknown) product reach a referral rate in the same order of magnitude? That would tell you a lot about the perceived quality of your solution.

Note: *Also interesting is the last line: this means, that on average, if you have a referral program, you can lower your Customer Acquisition Cost with 87% for each customer! This is what your experiment is for: how can you achieve a high referral rate?*

Still, it is difficult to find accurate benchmarks for every industry. These numbers may or may not translate to your situation. You'll probably have to try first, and then see if you can design an experiment to raise referral rates.

Step 2. Run the experiment
While running, keep track of who referred whom. Try to find your top referring customers. They are your biggest fans. Try to get in contact with them and understand what they love so much about your product. Also, try to find people that did not refer. Why not?

Step 3. Interpret the results
After running the experiment, get your data together and compare to your hypothesis.

[11.21] Recipe 21: Retention

About this experiment

When you're working towards Product Market Fit, retention is one of the most important ingredients in your Pirate Funnel, next to referrals.

If you get people to continue to use your product, if it is a subscription or if it is a repeat purchase, you raise their Customer Lifetime Value, but not their Customer Acquisition Cost. This will enable you to grow faster, because your churn is lower: you won't need to find so many new customers to replace customers that left.

A retention experiment is focused on increasing retention, and lowering churn. To do that, you first need to measure retention. Some metrics you might use are:

- In a subscription model: what is your (monthly or yearly) churn? How many customers leave?
- In an online model, how many daily active users do you have? How many people come back to your site or app?
- How many customers make a repeat purchase?

> By increasing customer retention by just 5%, a company's profitability will increase by an average of 75%.

Besides measuring, you need a way to influence retention and/or repeat purchases. Both of these metrics are things that will take time to measure. What are you changing that may influence retention and repeat purchases?

Retention:

- Changing pricing plans and giving discount when signing up for a longer period
- Keeping customers engaged
- Keep track of customers that have stopped engaging with the product
- Adding switching costs (like 'your data is in this app')
- ...

Repeat Purchases:

- Discounts for existing customers
- Drip campaigns, keeping people engaged
- Adding new products that are similar to what people already bought
- Re-engaging customers that bought products in the past (welcome back bonus)
- ...

Datasheet

Experiment: Retention
Type of experiment: Quantitative
Benefits:

- Learn how to increase retention and thereby Customer Lifetime Value

Dangers:

- This experiment can be long-running (it takes a while to see results)

Use this experiment to: Learn how to increase retention and thereby Customer Lifetime Value
Time and resources required to set up: Quite minimal, mostly find a way to keep better metrics
Evidence level: High, you'll see increase in repeat sales or lower churn. People commit.
Method: Create an offering for existing members or customers that increases repeat sales or lowers churn
Prototype: Your special offer to existing members or customers

Step-by-step

Step 1. Define your experiment
Design your experiment using appropriate benchmark numbers.

Churn
For SAAS companies, a 'good' churn rate of 5% (annually) is often mentioned. Research shows, however, that even many successful SAAS companies operate in the range of 7-12% churn annually. The medium annual churn in one study was 10%, or 0.87% per month.

You should calculate if your subscription model works for this churn rate first, before taking it as a benchmark. (You can use the Startup Calculator to quickly see different scenarios for different churn rates).

Repeat sales
A good benchmark for a business is to have about one-quarter of customers return to repeat a sale. For e-commerce companies, however, repeat sales account for as much as 50% of total sales. (Source: Investopedia)

***Note:** This is for established companies and brands, you may set the bar a bit lower when you're just starting. For WRKSHP, for instance, it is around 25% at the moment.*

Step 2. Create the prototype
Come up with (a few options for) an offering towards existing customers or subscribers that you think will increase repeat sales or lowers churn.

Test your offering with a few test-subjects (chosen from your existing customers). Talk to them about this idea and see how they respond. Adjust if necessary.

Step 3. Implement and run the experiment
Keep track of the key metrics while you are running the experiment. As it will probably (especially in the case of churn) take a while to see results, try to identify if you are running other experiments in the time frame that may affect the results.

Step 4. Interpret the results
Use your data to see if you can validate your assumption.

[11.22] Recipe 22: Pricing

About this experiment

It is key to know what a good price is for your product. This experiment is a simple experiment you can use to see if you have a good price.

The idea is to:
1) Present different customer groups with a different price, and comparing the results.
2) Present customers with different price packages, and comparing the results.

You can show different prices and packages to different customers at the same time, or you can change the price over time (see Recipe #23, Raise the Price).

Another thing you can do is keep your prices the same, but change what is in each package. Move features between packages and see which features have an impact on conversion rates.

For the different situations, you'll measure the conversion rates of the different packages. How many people buy your product? What is their favourite package?

Datasheet

Experiment: Pricing
Type of experiment: Quantitative
Benefits:

- Find out what your customers find an acceptable price for your product.

Dangers:

- You'll be checking multiple options, and that means you'll need a large number of samples!
- If you do this too early you may get the price wrong, value can change over time

Use this experiment to: Find out what your customers find an acceptable price for your product.
Time and resources required to set up: Not very great (unless moving features between packages impacts implementation)
Evidence level: High
Method: Create 3 different pricing packages with 3 different prices each, and measure conversion
Prototype: bla

Step-by-step

Step 1. Define the experiment

Design the experiment, using a suitable conversion rate benchmark from either your experience or competing products in the market.

Step 2. Define the pricing

Create three packages, ‘light’, ‘medium’, and ‘large’, with three feature sets.

Feature	Package 1 (Light)	Package 2 (Medium)	Package 3 (Large)
Basic feature 1	X	X	X
Basic feature 2	X	X	X
...			
Medium feature 1		X	X
Medium feature 2		X	X
...			
Advanced feature 1			X
Advanced feature 2			X
...			

Your ‘Light’ package should be the absolute minimum functionality that is useful to people. ‘Medium’ is a good mix, and ‘Large’ is everything and the kitchen sink. You would like to find out how the ‘Light’ and ‘Medium’ packages are doing, you’d rather not have to actually build the ‘Large’ package, it has too many features.

Come up with a basic price for each plan.

10X Rule
To do that, first look at your different customer segments. Look at how much the solution you offer is worth to them.

- What costs are they **saving** because they can use your solution? (Does it replace a more expensive product?)
- How much **time** are they saving?
- What **opportunities** do they gain because they use your product?

Use this to calculate a value for your customer.

The 10X Rule says people should be able to get a 10X ROI from your product. If for instance a freelancer is able to save €100 per month from your product then your Freelancer plan could be priced at €10 per month. If enterprises are able to make €10k per month because of your product, price your Enterprise plan at €1k per month.

As a sanity check, get an idea of how much they spend currently on products such as yours. This is what their frame of reference will most likely be. If they are used to products that cost €10, they might balk at a price of €100, even if their 10X ROI is €1000.

Let's call this basic price 1X.

Define your basic price for each package.

Variant	Package 1 (Light)	Package 2 (Medium)	Package 3 (Large)
Basic Price (A)	0.5X	1X	3X

Now, create variations in pricing for each package. Note that package 3 is substantially overpriced. You are really measuring Package 2 for the most part.

Variant	Package 1 (Light)	Package 2 (Medium)	Package 3 (Large)
Basic Price (A)	0.5X	1X	3X
B	1X	2X	5X
C	2X	3X	7X

Option C is only there to make sure you're not underpricing.

Test this table with some test subjects from your customer segment. What do they think?

Step 3. Implement and run the experiment
Implement the pricing tables, and run the experiment. Keep track of the different conversion rates for each combination.

Step 4. Interpret the results
Interpret the results and compare with your hypothesis.

Hacks

Besides price, you could also experiment with different *value metrics.* The ***value metric*** is the basis on which you charge your users. For instance, BrowserStack charges its customers on a **per user** basis, Stripe charges on a **per transaction** basis, and Mailchimp charges on the basis of **number of subscribers**. If you do this, make sure your customers understand your value metric.

[11.23] Recipe 23: Raise the Price

About this experiment

This experiment does exactly what it says it does. Raising the price incrementally is nifty way to figure out if there is interest and getting pricing signals at the same time. The first person to buy the product or sign up pays the lowest price, the second a slightly higher price, and so on. You keep raising the price until it starts to become hard to convince people to buy.

In this way, you start out making easier sales, and with each sale you make that a bit harder, forcing you to learn what is really important in the sales process.

This can work online, but it is easier to do when you sell face to face, in a B2B setting for instance. If you do it online, you can see that at a certain point a higher price means that conversion rates drop, but you probably won't find out what exactly was the reason for not buying.

Example
For the book Business Model Generation, Alex Osterwalder and his co-authors used this approach to raise awareness and interest people to become a proofreader and ambassador.

Datasheet

Experiment: Raise the Price
Type of experiment: Quantitative
Benefits:

- Start with easy sales and force yourself to learn with each sale
- Don't hit a wall when you start with a price that is too high

Dangers:

- Online, you'll need to be consistent in the price hikes, and explain to people when and why you do it, in order to avoid frustrated customers. In a situation where people don't see a list price (e.g. in many B2B situations) this is less of a problem.

Use this experiment to: Find out your price point and learn about your sales process and messaging
Time and resources required to set up: You'll need something to (pre)sell. Otherwise, it's very lightweight.
Evidence level: High. People will commit to something and they will hopefully also tell you why they don't want to buy.
Method: Raise your prices every few sales.
Prototype: A product to sell and an easily changed pricing

Step-by-step

Step 1. Have something to sell
Refer to e.g. Landing Page, Pre-Sell, Pop-up, and MVP recipes. You'll need to have (or fake) something people can buy.

Using that, define an experiment. Coming up with a benchmark is a bit more difficult here, because the conversion will definitely be impacted by a changing price.

Define a starting price (which can be quite ridiculously low if the product doesn't cost you too much out of pocket). Also define how and when you will raise the price.

Step 2. Start selling
Get people to see and buy the product. Sell a certain number of products in a time frame. Calculate the conversion rate (visitors / buyers, or prospects / buyers). If possible, use the sample size calculator to come up with a low bound on the number of sales.

Step 3. Raise the price
Raise the price by the specified amount (even if you're selling way more, don't go up too fast). Again, wait for a low bound of conversions, and iterate. Plot your price / conversion values in a chart. Continue raising the price until you see the conversion start to go down. Then raise it some more. Once your conversion really dips, you've reached a price point.

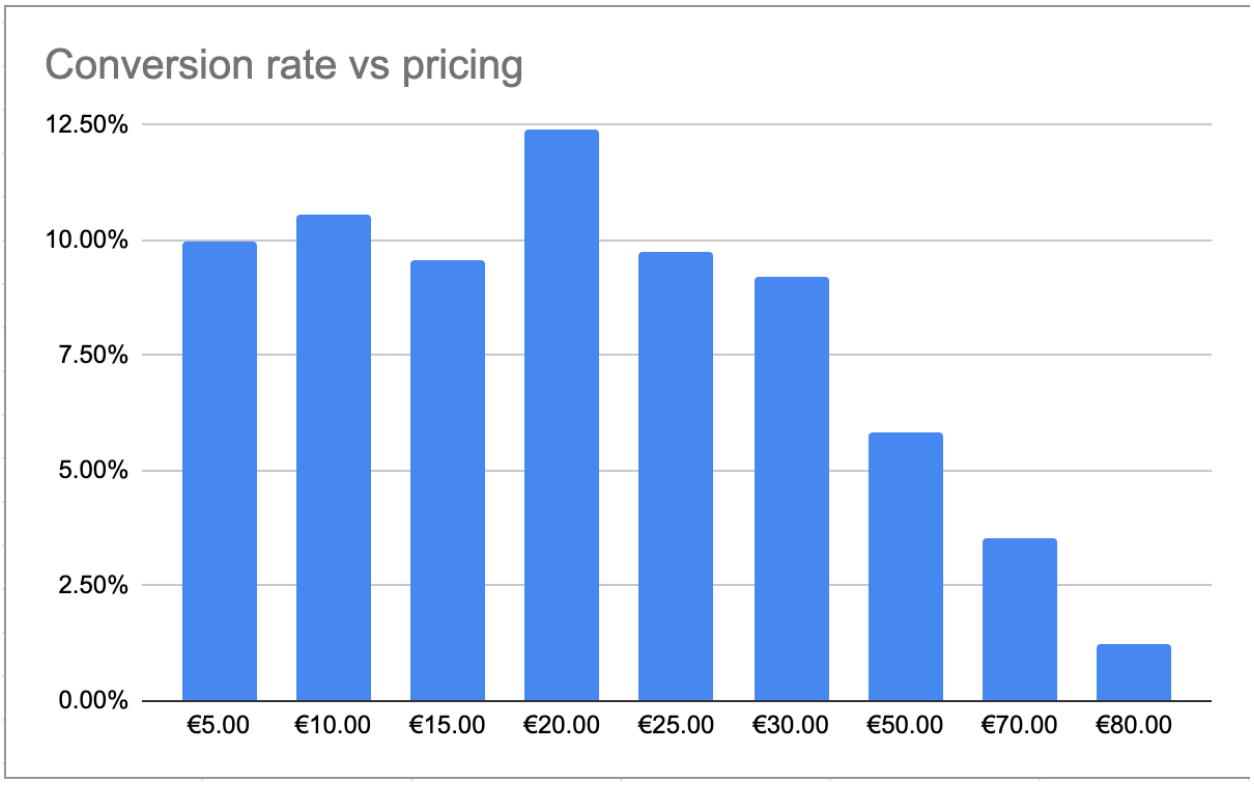
Step 4. Calculate the optimum
Using your chart, find the spot where prospects x conversion x price is optimal. Check if that is above the level you need to make a decent profit.

Example

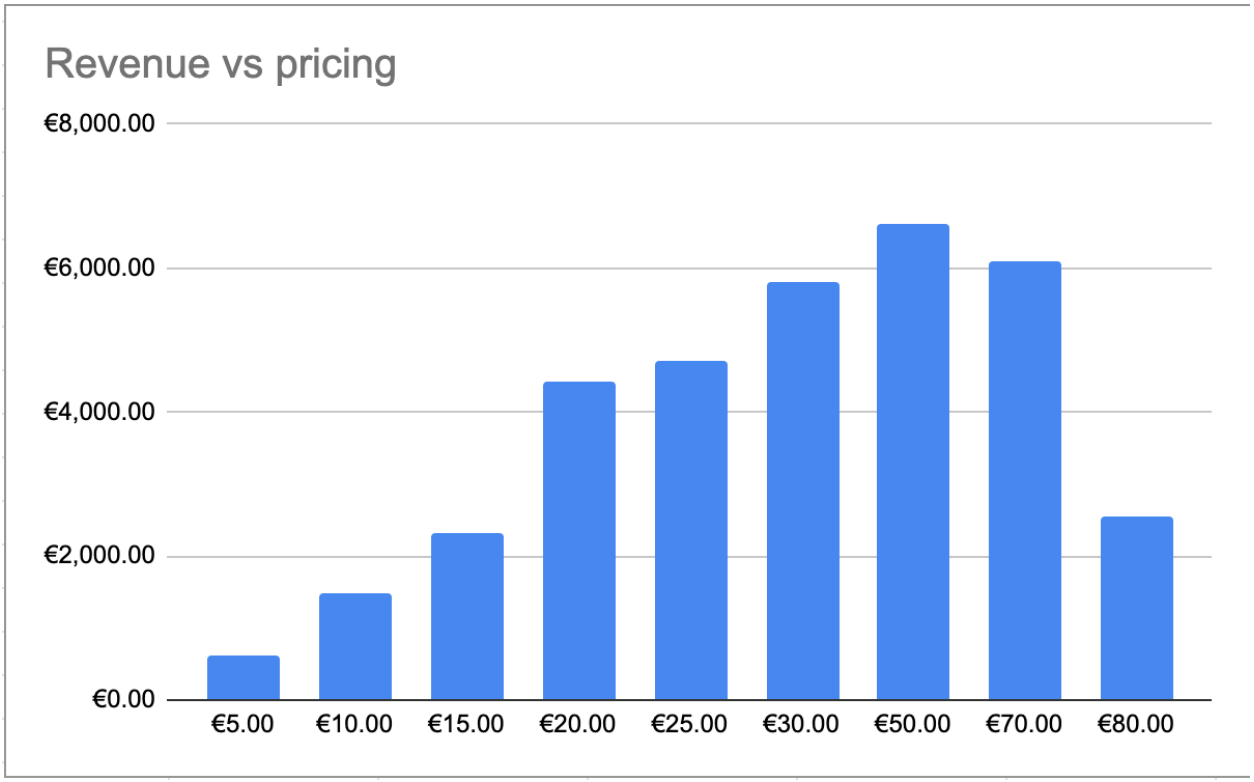
Say you have the following data from your experiment:

Price	Prospects	Growth Prospec	Sales	Conversion	Revenue
€5.00	1213	100.00%	121	9.98%	€605.00
€10.00	1409	16.16%	149	10.57%	€1,490.00
€15.00	1620	14.98%	155	9.57%	€2,325.00
€20.00	1784	10.12%	221	12.39%	€4,420.00
€25.00	1939	8.69%	189	9.75%	€4,725.00
€30.00	2102	8.41%	194	9.23%	€5,820.00
€50.00	2261	7.56%	132	5.84%	€6,600.00
€70.00	2470	9.24%	87	3.52%	€6,090.00
€80.00	2631	6.52%	32	1.22%	€2,560.00

Plotting the conversion rates in Google Sheets show you that conversions are dropping above ~ €30. The fact that the conversion is lower at the start is probably because you needed to create awareness and lead people to the product first. The number of visitors in this example keeps growing, as more and more people become aware.



This is interesting information, but what you really want to know is where the revenue starts to drop.



Graphing Revenue vs Pricing gives a clear picture. The optimal price seems to be around €50.

Resources

The example above can be viewed here as a Google Sheet.

[11.24] Recipe 24: Beta Test / Soft Launch

About this experiment

A Beta test an early attempt to test that you're solving the basic customer needs which you have identified, your solution solves your customers' needs in an acceptable (if basic) way, and that you can acquire and activate customers.

Note: *It's important to recognise you still have a lot of learning to do. A Beta release is not an excuse to get away with a sub-standard products or a shortcut on quality, it's a controlled experiment to maximize learning before scaling.*

Consider this activity as an opportunity to learn and maximize value, while at the same time reducing the risk of discovering damaging issues after launch. A Beta test also gives you an early opportunity to talk to users that are actually using and experiencing your product.

There are private (invite only) beta tests and public (open) beta test. The former is more controlled and reduces the risk of reputation damage more if your beta turns out to have issues. The public beta test is something you can use to transition to a full MVP after you have done a private beta.

In a private beta, because you limit the number of users to the ones you have invited, you're able to see how users respond before you have tackled (operational) scaling and security problems that might occur if you had thousands of users. Typically, solving these problems can be expensive, and it is preferable to see if the features you have implemented are satisfactory.

Besides a full MVP, running a Beta test requires the most work of all the recipes in this cookbook. You'll need to have a functioning product, have the onboarding and support in place, and you'll have actual users.

Datasheet

Experiment: Beta test
Type of experiment: Mixed
Benefits:

- Observe how people use the product
- Talk to early users
- Start learning about scaling (the operational side of) your product

Dangers:

- If the beta is (perceived to be) crappy by customers, this can lead to reputation damage.

Use this experiment to: Learn about how people use the product and get feedback from early users
Time and resources required to set up: It takes a lot of work, since beta features need to actually work. No more 'fake it before you make it'.
Evidence level: High. It's quite close to the actual product, and people are really using the product.
Method: Create a first beta test version with core functionality and track people using it
Prototype: A first beta test version with core functionality

Step-by-step

Step 1. Define your experiment
Design your experiment, by first coming up with suitable riskiest assumptions to test. The Beta test is usually something you do when you are past problem solution fit, and it is most useful in finding out how users respond to the features of your product.

- Do they enjoy using the product?
- Do they come back?
- Do they want to refer it to others?
- Is the product 'good enough' to be acceptable?
- Do people really use feature X?
- How often do people use the product?
- ...

Assumptions that are in line with these questions can be tested with a beta test.

Step 2. Prepare the experiment
Before the experiment, try to find out from invited users what they expect from the product. What are they excited about? Usually, this can be done as a survey, that needs to be filled in in order to get access to the beta.

After the experiment, you'll want to debrief beta users in (preferably) an interview or a survey. You need to understand what they did, and what they liked or hated.

Tip: *if you're inviting beta users, tell them you want to schedule time with them for a debriefing interview as part of the test. Just sending out a survey will mean that people that don't like the product or find fault with it might skip sending you feedback.*

Step 3. Run the experiment
Make sure you are able to monitor what people are doing with your product. With digital products, this can be done using event tracking. With physical products, you'll need to be creative.

One idea is to ask people in the beta to take smartphone pictures of how they are using your physical product, and upload them to you, or to keep a small diary. You might prompt them once a week through a ‘check in’ email or a survey to see how they are doing.

Step 4. Debrief
Debrief users after the experiment ends, to get rich feedback .What did they really like? Or hate?

The most important question to ask at this point is:

“What would your response be if you could never use this product again?”

The answer should be in the range of [Happy | Slightly Happy | Don’t Care | Disappointed | Very Disappointed]. This will give you a good idea of how important your solution has become to them.

Step 5. Interpret the results
Gather the data and make a decision on how to continue. Was your initial question answered? What else did you learn that was not expected?

Hacks

You can learn about scaling as you go, by starting with a private beta with a limited number of users, and adding new users in batches. This combines well with a waiting list.

[11.25] Recipe 25: MVP Test

About this experiment

For this experiment, you're building and launching a Minimum Viable Product (MVP).

> The minimum viable product is that version of a new product a team uses to collect the maximum amount of validated learning about customers with the least effort.
>
> -- Eric Ries

As you can see in this definition, there is a lot of room to fit your MVP to what you want to learn about your customers. You could use an MVP approach to test specific features, your sales process, your onboarding, and other things.

And, the term MVP has been used for all these examples, and many more. Some people call a paper prototype an MVP, and others reserve the term for a finished product. We need a more precise definition.

This experiment uses an approach where the MVP is a **'complete' product** that can be used by **users you may not know or be in direct contact with** (as opposed to the private beta test, recipe #24). What that means can be defined by looking at the words Minimum Viable Prototype.

Minimum
The word minimum tells you that this is not going to be a bloated product with thousands of features. It should be minimalistic and focused. What should it be focused on? That is where the other two words come in.

Viable
It is called a minimum **viable** product. The user needs to be able to experience the product as if it is a real product. It may be minimal, but it has functionality. For it to be **viable** in the eyes of the customer, it needs to **solve their problem**. It's more than just a few cosmetic features put together to see how people react.

Product
The word **product** also tells you that the prototype is to be experienced as something that can be experienced as a **'complete' product,** by **users you may not know or be in direct contact**

with. It should be a complete enough experience to see how users respond to the entire user experience from start to finish, **without your help or guidance**.

This does not mean you need to handhold your users too much and build elaborate tutorials (one of the goals could be to learn how much hand holding they need), but you will need a minimum of instruction to make sure they can work with the product by themselves.

Core functionality
Combined, this means you should focus on the **absolute core functionality**. What is the one function of your product that it can't live without? And what are the (few!) extra supporting functionalities that are needed to make it a complete user experience? Forget about the nice-to-haves, the functionality that gets used 1% of the time, and anything else you can get away with. Only focus on giving the user a new way to solve their problem.

This is not a drill
You are selling or opening your MVP to real users out in the real world. This means you're no longer measuring responses of people to a hypothetical product or a test version, they are using your product for real. They want to solve their need with it. If they are not positive about your core functionality or it does not do what they need it to, that means they will, in the best case, abandon it, and in the worst case, leave negative reviews online.

If you're running an MVP, you are making a commitment to users that they will be able to use and access your product.

Datasheet

Experiment: MVP test
Type of experiment: Quantitative
Benefits:

- You are testing with a real product, with real users, who make real commitments.

Dangers:

- The MVP is (usually) public. Negative feedback on an MVP can negatively affect your reputation.
- Users respond to what they get, not what it might become. If the product is 'too minimal', they might not see the benefits.

Use this experiment to: Find out how people respond to the core functionality of the product.
Time and resources required to set up: This may take a lot of time, because you need to have a functional, viable product
Evidence level: High.
Method: Launch a minimal version of the full product, get responses.
Prototype: MVP

Step-by-step

Step 1. Define your experiment
First, find out what you want to learn. What are your assumptions, based on all your earlier experiments, of how users will use the product? Are you interested in how they use certain features? How often they use the product? The retention rate? The referral rate? Come up with a hypothesis based on this.

Step 2. Run the experiment
Prepare a launch for your MVP. You can do this as a soft launch (without using big publicity) or you can go all-in and have a launch party, do marketing, and get PR involved, which is riskier. In any case, you'll need to come up with a strategy to get people to learn about your MVP, sign up, use it, and pay for it. Look at the Pirate Metrics Canvas and come up with an approach for each of the steps in the funnel.

Setup and start your marketing campaign and get your first signups. Start onboarding users, and monitor how they use the product.

If possible, try to contact users and get their feedback.

Step 3. Interpret the results
When the time you set for the experiment has passed, interpret the results. Were your questions answered? Was your product viable enough? What do you need to change?

You can use continuous improvements to keep your MVP running and add the features you conclude people are missing.

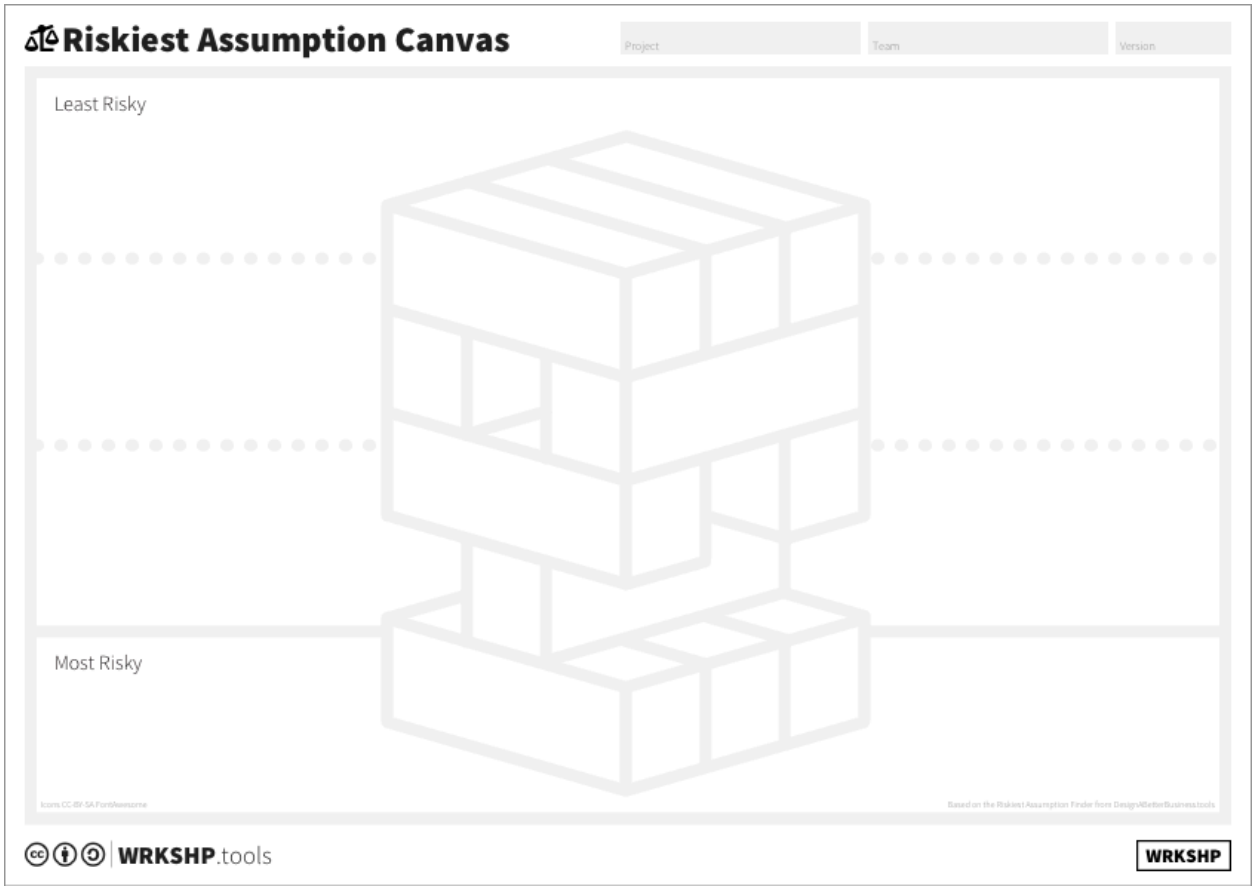
[12] Appendix: Tools

[12.1] Tool: The Riskiest Assumption Canvas

- Materials: Flip over paper, coloured post-its, markers
- Time: 20-30 minutes
- Difficulty: Easy
- Number of people: 3-5

We've all been there: your idea is so great that you're literally bursting at the seams wanting to launch it as soon as you can (maybe even today!). Most of us feed on this excitement. But how do you know you're making the right bet with your idea? Which bets does the success of your idea hinge on? These are your riskiest assumptions; they need to be tested.

Whether you work for a small start-up or an existing large organisation, validate your riskiest assumption as quickly and cheaply as possible so you don't waste valuable time and resources toiling away at something that likely will never work. But this is often harder than it sound. How to find your riskiest assumption?



Canvas Building Blocks

The Riskiest Assumption Canvas has four main areas, organized from top to bottom. The 'Jenga® tower graphic' in the center is not part of the building blocks, but only serves to illustrate that the bottom area is more risky than the top.

The topmost horizontal box is reserved for the least risky assumptions: the ones you are most certain of. This is also where (in)validated assumptions end up: validating means reducing risk.

The bottom most box is for the single most risky assumption.

In between, there is space for assumptions that have some risk, but are neither the riskiest or the least risky.

How to use the Riskiest Assumption Canvas

In your stack of assumptions, the riskiest one is the first gate. If when testing this assumption it continually comes back as “false,” you don’t get to pass go, you don’t get to collect your $200. This tool will help you rank your assumptions before moving on to experimentation.

The first key to identifying your list of riskiest assumptions is to bring a team together to unpack the idea and brainstorm together.

To identify the assumptions you have made about your business, have a look at other canvases you made, such as your Why, Design Criteria, Customer Journey, or Context Canvas. Each of these canvases contains untested assumptions.

Especially the Business Model Canvas (see below) makes it easy to find assumptions, by looking at each of the building blocks and asking yourself: is this really true? The goal of this exercise is to sort these assumptions, and find the one riskiest assumption that underpins your idea. The one assumption that, if it is not valid, will make your idea come down like a Jenga® tower.

With your team in place, use your designer tools (sticky notes, markers, and a big wall) to rank these, based on which you could not do without or which are most likely to be false. The sooner you find these, the more likely you are to be able to validate them and either move forward or pivot.

Finding the riskiest assumptions

Finding the riskiest assumption is not always easy. Discussing assumptions with your team will help to identify the ones to go after. Do this visually so it’s to the point and provides you with the outcome you need! This tool makes it easy for your team to have meaningful discussions on what the riskiest assumption really is, and provides a way to come back to the discussion after validating.

Jenga®

Jenga® is a game where players in turn try to remove blocks from a wooden tower. Each block that’s pulled out may make the tower collapse, but the blocks on the bottom are critical to keeping the tower upright.

Think of your idea as a big tower, where all of the bricks are assumptions. When one of the assumptions on the bottom of the stack is invalidated, and the brick is removed, the entire tower may fall. When you remove one from the top, not much will happen.

To make your idea succeed, you need to make sure that the base of the tower is safe. We need to start at the bottom, with what we call the riskiest assumptions. At the moment, All the other assumptions are not as important. After all, if the riskiest assumption is incorrect, it may be totally irrelevant to think about any of the others: maybe your idea needs to change completely in the light of the new knowledge!

The goal is to try to make the tower fail fast! So check the bottom-most assumption first, which is the riskiest one. That's what you'll want to know more about. If it is right, you can move to the next riskiest assumption. But if it fails, your Jenga® tower falls, and you'll need to go back to the drawing board to find another approach that works better.

Step 1. Identify Assumptions

What are your assumptions? What are the things you're not sure about? With your team, start by just writing all your assumptions on sticky notes, but don't stick them on yet. Refer to the war room and your point of view for inspiration.

Step 2. Map your Assumptions

Then, put the assumptions onto the template, each team member placing them in the middle 3 boxes, where they think it's best. Don't discuss yet!

Step 3. Arrange the Assumptions

Once all of the post-its are on the wall, each team member quickly goes over his or her post-its, and explains why they think they should be in the place they are.

Now, with your team, take turns moving sticky notes around. Try to find out which assumption is the riskiest one. Go over each post it, and see if everyone agrees that it is in the right spot. Remember: the assumptions that absolutely must be true for your idea to work go on the bottom of the stack.

The ones that are less important or depend on other assumptions go higher up. Compare pairs of post-its to see which one has priority. When sticky notes move back and forth between boxes, put them halfway between.

Step 4. Fundamental Assumptions

Finally, go over each box and see if there are any assumptions in there that really depend on others (move them up) or that are fundamental (move them down).

After about 15 minutes, you should have only a few left in the lowest box. Vote with your team as to which one you think is the most fundamental one.

Step 5. Describe your Riskiest Assumption

Now that you have identified a riskiest assumption, try to describe what it means with your team. When is the assumption true? Does it depend on something else? How can you make it s.m.a.r.t.?

Step 6. Next Steps

Next, design an experiment using the Experiment Canvas

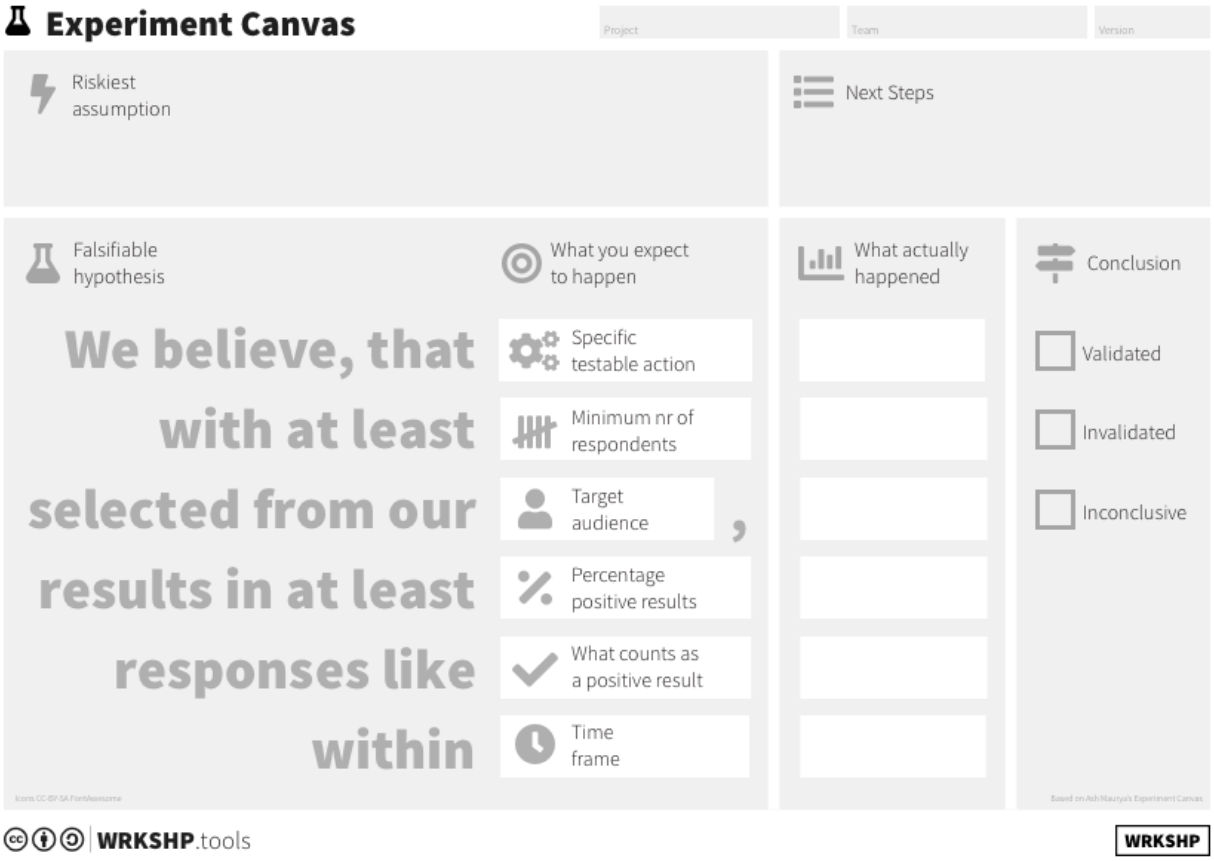
Checklist

✅ You clearly identified **one** riskiest assumption
✅ You have described your riskiest assumption

[12.2] Tool: Experiment Canvas

- **Materials**: Flip over paper, coloured post-its, markers
- **Time**: 20-30 minutes
- **Difficulty**: Medium
- **Number of people**: 1-5

Once you've found your riskiest assumptions you'll need a way to figure out how best to test and measure them in a quantitative way. The experiment canvas, created by Ash Maurya, provides a straightforward way to break down your assumptions into measurable, observable, experiments.



Canvas Building Blocks

- Riskiest Assumption: the assumption you want to validate

- Falsifiable Hypothesis: a statement defining your hypothesis, the way you will know your assumption is valid or invalid.
- What actually happened: the results for each of the items in the hypothesis
- Conclusion: the outcome of the experiment
- Next steps: the next steps you take (pivot or persevere)

How to use the Experiment Canvas

The purpose of the experiment canvas is to design the right experiment at the right time, facilitating a team to have the right conversation. With the experiment canvas, it is easy to design a well-defined experiment

Track the data immediately and write everything down, so that later you can check if you interpreted the results correctly.

Step 1. Select the Riskiest Assumption

Start with identifying the current Riskiest Assumption and describing it in a way that you want to test. What will it mean for your idea if this fails? How can you tell?

Step 2. Define your Hypothesis

Specify a clear, falsifiable hypothesis and experiment setup. After running the experiment, check the results and plan your next steps.

Your hypothesis is a statement you believe to be true about your riskiest assumption. Write it down before you run the experiment. It is too easy to change the conditions afterward to make the data fit, and this robs you of valuable insight.

For this version of the canvas, the original hypothesis formula is extended so it is easier to make it quantifiable:

"We believe, that (specific testable action) with at least (minimum number of respondents) selected from our (target audience) results in at least (percentage) responses like (what counts as a positive result) within (time frame)"

Each of the elements in brackets need to be quantified.

It’s okay to use a bandwidth for this, as long as you specify it upfront. The metrics you define need to be actionable (i.e., they need to directly relate to the hypothesis) and accessible (i.e., you need to be able to see the results).

Try to find benchmarks to define percentages, and allow for the fact that if you have small numbers of respondents, you will need quite large measurements to be sure. Look for large percentages.

Tip: Link the numbers back to the assumption you are testing. Why does having 10 positive results validate your assumption?

- **Specific testable action:** This is what you will do as your experiment. It usually ties in with the prototype and the method of your experiment. Examples are: get people to click on an online ad, run interviews, or have a user test.
- **Minimum number of respondents:** You will need some lower bound on this. If you fail to meet that lower bound, you won't be able to get any results from your experiment. If you have an experiment with a clear yes/no answer, go for at least 30-50 people. More is better.
- **Target audience:** The group of people you will select from. Also think about how you will select them. Is it random? Or do they already know you?
- **Percentage positive results:** The percentage above which you will define the experiment outcome as validating your assumption. When talking to people, try to go for larger percentages (>50%). If you have an online experiment, try to find a conversion benchmark, and aim for at least double.
- **What defines a positive result:** The type of answer or customer behaviour you count as a positive result. In the best experiments, it is a commitment the customer gives. This can be a small commitment, such as giving their email address or showing up at a meeting, or a larger one such as giving access to network, or actually buying a product or service. The stronger the commitment, the more significant the result. This aspect is linked to the protocol you define.

Armed with this hypothesis you're ready to start your experiment.

Step 3. Define a Protocol

Having a good protocol to run the experiment is key to quality results. It's a complete science to do this right, but for our purposes it is sufficient to just avoid the biggest mistakes.

- Create a Google Sheet accessible to all the team members that will run the experiment.
- Create a uniform list of questions or observations. Include some questions to establish rapport with your respondent.
- Define how you will demonstrate prototypes, and ask questions. How long will the experiment take?

- Define how you will score the results. E.g. when your respondent volunteers they have a certain problem, that counts as 5 points. When they only acknowledge it when you ask about it, it's just 1 point. For a prototype, the amount of time the respondent is engaged could be part of the score, or the number of features they tried out.
- Make sure your scores are meaningful.
- Define how you select respondents
- Define how you'll tally up the score, and what score will mean the interview or demonstration counts as positive. Specify any qualitative outcomes and variables. What different answers you are expecting? How will you cluster them?

Step 4. Build the Experiment

Here, you'll need to come up with the questions or prototypes you need in detail.

The materials you'll create fall broadly in three categories: interviews, offline prototypes, and online prototypes.

Have a look at the prototyping part of this module for more details.

Note: At this stage it is rare that you need to test a technical prototype. In most cases, if it actually works in reality only becomes interesting after you have made sure people are waiting for it. So when we say '*prototype*' we really mean something that looks just real enough that customers are able to react to it in a meaningful way. And when we say 'just real enough' that is really telling you to do the bare minimum. No difficult branding exercises, technical setups, or scalable solutions!

Never forget the golden rule:
KEEP IT SIMPLE!

Step 5. Run the Experiment and Collect Data

Ok, so you have an experiment setup. Now it's time to get out there and collect the data. Use the protocol to your advantage! Make sure you record everything of value (and use for instance a google sheet to fill in the data).

Step 6. Interpret the Data

Once all the data is in, it's time to go over it. Get your team together and score the data. For qualitative experiments, have a look at the Experiment Outcome Canvas also included.

Step 7. Draw Conclusions

Once you have your data all interpreted and scored, it's time for conclusions. Get your scores together and compare what you got with the hypothesis you setup.

1. Do you have enough respondents?
2. Are the respondents in the right target group?
3. Are the respondents diverse enough?
4. Did you get enough positive scores?

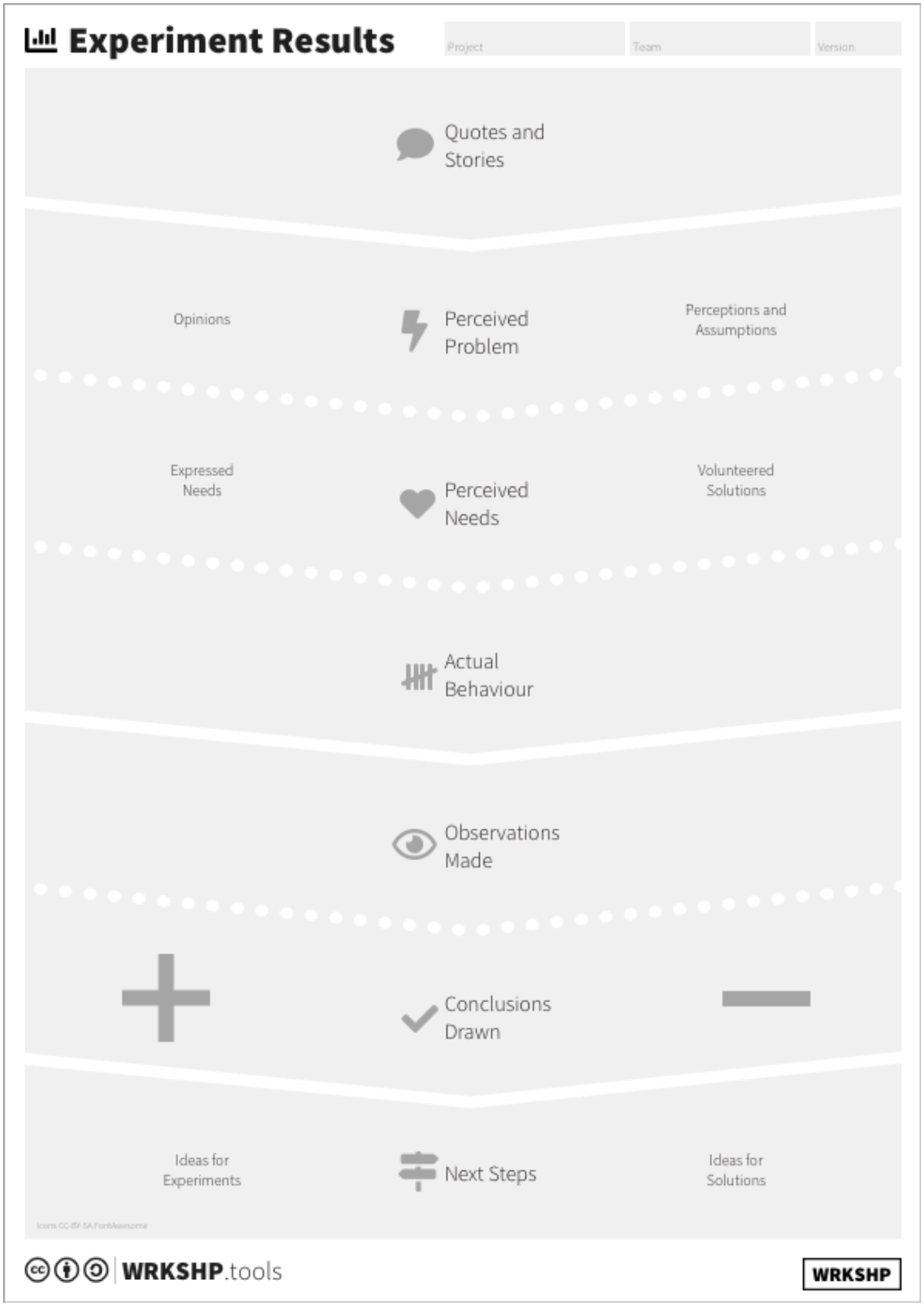
If you can answer 'yes' to all of them, then you'll have a validated assumption. If you can answer 'yes' to points 1-3, but didn't get enough positive scores (point 4), you'll have invalidated the assumption. If you can't answer points 1-3 with 'yes', you'll likely have a botched experiment, and the result is inconclusive.

Step 8. Next Steps: Pivot, Redo, or Persevere

Ok, you've got your conclusions. Time to act on them. There are three paths forward:

- **Validated**: You can persevere. Pick your next Riskiest Assumption and do an experiment to validate it.
- **Invalidated**: You'll likely need to pivot. Go back to the drawing board and see how you can modify your idea armed with this new knowledge about your Riskiest Assumption.
- **Inconclusive**: You'll need to take a hard look at your experiment. Did you perform it right? Did you have the right setup? The right hypothesis?

[12.3] Tool: Experiment Outcome Canvas



Step by step

The feedback you received from the interviews is split in four big categories (from top to bottom):

- Quotes and Stories
- Perceived problem, perceived needs, and behaviour
- Your observations and conclusions
- Next steps

1) Quotes and Stories

The first category is filled with raw quotes and stories selected from interviews.

2) Perceived problem, perceived needs, and behaviour

The second category splits results in information about the respondent's perceived problem (how they experience the problem you want to solve), their perceived needs (what they tell you about what they think they need), and their actual behaviour (what they have already done in the past to deal with the problem).

This distinction is important, because it is so easy to pick up only on what you'd like the respondent to answer to your question. It's so easy to hear that they like your solution, or that they really need it. But that information is close to worthless. (They're probably lying — or being polite).

Solutions and opinions volunteered by respondents, telling you how they might solve the problem in the future, are also close to worthless information. People don't know what they will or won't do in the future, and they have a very difficult time predicting their own feelings.

The thing to look for is behaviour. Have they actually experienced the problem in the past, and did it bother them enough that they actually found or tried to find a solution for it? That is what you need to hear if you're looking for information coming from early interviews. It's much harder to 'be polite' about actual behaviour. It's the actions that count, not the words and opinions.

Sticking this information in separate boxes means it is all there, but it's organized in an 'evidence pecking order'. The behaviour box is the most important one. But, if a lot of people say the same things, or have the same opinions, you might want to run a separate experiment based on that and see if their actions reflect those opinions.

3) Your observations and conclusions

The third category can be filled with the interviewer's notes. What observations were made by the interviewer? What conclusions did they draw?

It is important to keep this information separate so that it won't get mixed with the results coming from the respondents.

It's great to collect these observations, and they may help you a lot, but they are **your** observations. They are something that you added — and therefore, based on what you already knew **before** you conducted that interview. They reflect your view of the world and your biases more than anything else.

4) Next steps

Finally, there is space for next steps. What follow up questions would you like to ask? What other things would you like to know?

This tool can be used as a guideline to structure unstructured interview results, and to cluster results together. You can use it to prepare for the workshop and pre-process your interview results. In that way, you can use your time in the workshop to make sense of the results, and, once you have done that, vote—but in an effective way.

Now at least, if you do place your dot vote sticker next to an interesting behaviour, it'll be clear that that is much more valuable than a vote for your own observation or opinion.

You can more or less see from where the dots are on this canvas if they are more likely to be influenced by confirmation bias.

Finally, using this tool can help to compare results over time coming from different interview settings with different formats.

[12.4] Tool: Pirate Metrics Canvas

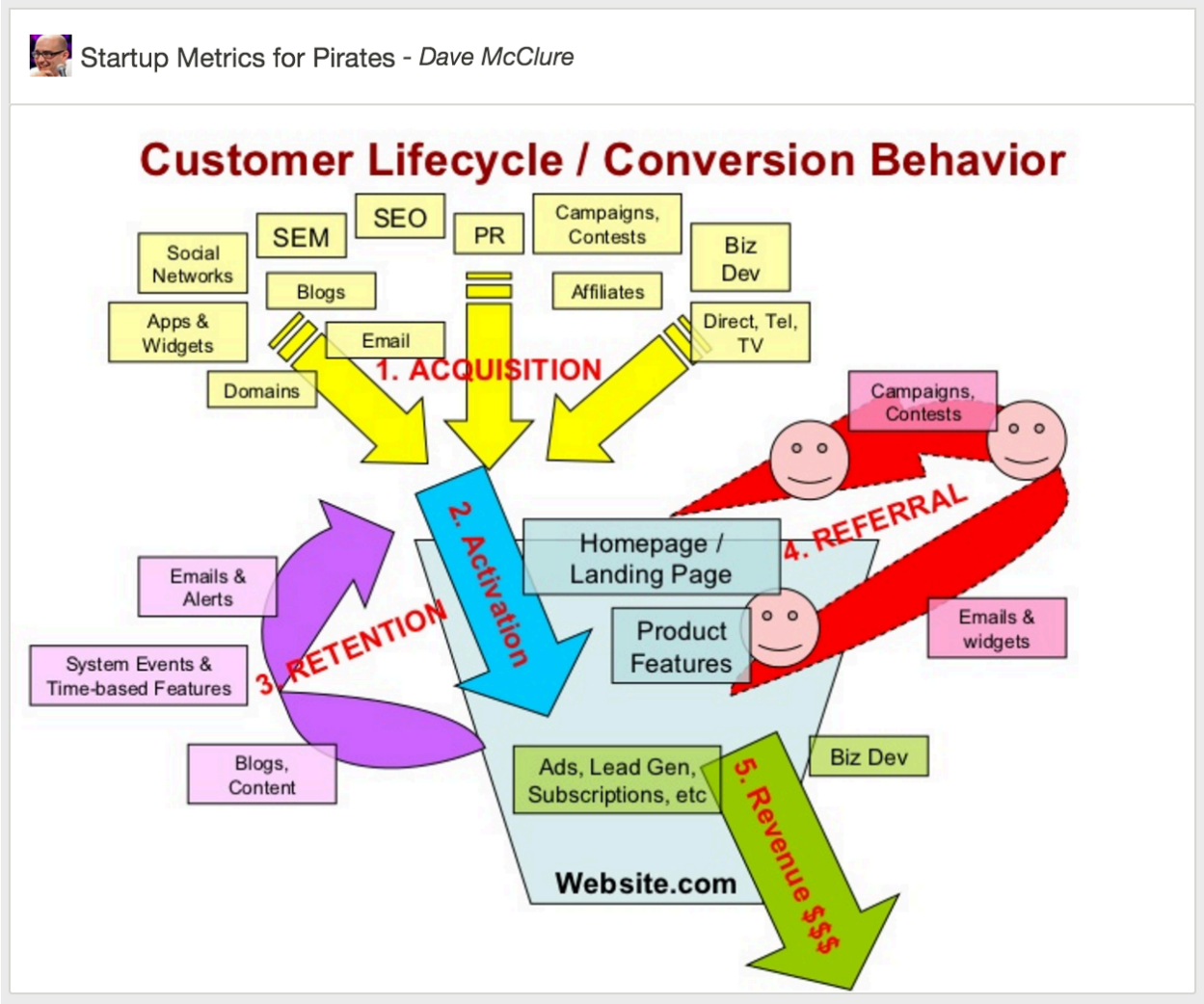
When you are developing a new digital service or startup, it is vital to understand how you will reach your customers.

And to do that, according to Dave McClure, it's best to think like a pirate. Dave McClure came up with Pirate Metrics and first presented it in 2007 at Ignite Seattle.

Video: https://www.youtube.com/watch?v=irjgfW0BIrw&feature=youtu.be

The idea is to map out the entire journey a customer makes to go from someone who doesn't know the service at all through different stages to someone that is a paying customer. What are the different stages a customer goes through? What are the decision points? What are the conversion rates between subsequent stages?

Knowing this will help you calculate those all important metrics: the CAC (Cost of Acquiring a Customer) and the CLTV (Customer Lifetime Value). If you know your CAC and CLTV, you also know what your profit margin is.



The actual model behind it looks quite complicated.

A customer has to go through these stages to go from someone that does not know about the service all the way to a customer that generates revenue. Each of the transitions between stages has a conversion attached to it. Not all customers that you acquire end up using the service, and not all of these may become paying customers. That means, you will lose customers going from stage to stage.

To get more paying customers, you can do two things. You can get more customers into the system (by working on the acquisition stage), and you can increase the conversion rates on each of the stage gates. Both strategies, typically, come with a cost. The total cost that you incur in order to get a customer from the acquisition phase to a paying customer is what determines your CAC.

Strategic Questions

Typically, this means early stage startups have questions such as:

1. What are the channels we have available to acquire customers?
2. What strategies can we use to influence the conversion rates between stages?
3. Where should we spend our money in order to lower CAC, raise CLTV, and increase user growth?
4. Should we focus on getting more active users? Or more paying customers?

Strategically, finding the answers to these questions is super important. That means that discussing the current state of the Pirate Funnel and the strategies to the future should be an integral part of the innovation journey, especially once you start to work towards product-market fit.

To enable that discussion and make it easier to keep track of the current situation and the focus for the near future it is super helpful to have a visual aid for team discussion that you can come back to check later: a canvas.

The Pirate Metrics Canvas

I have turned Dave McClure's idea into a canvas that you can use to make sense of what the Pirate Funnel looks like for your service or startup. To make discussion easier it is depicted in a linear way (from left to right).

Canvas Components

Pirate Metrics defines 5 conversion steps (grey):

- 1. Acquisition: People that are interested in the service.
- 2. Activation: People actually using the service
- 3. Retention: People that come back to use the service
- 4. Referral: People refer the service to others, helping in acquisition
- 5. Revenue: People that pay for the service

The canvas adds to this 6 main stages (white):

- Potential customers
- Acquired customers
- Active customers
- Retained customers
- Referred customers

- Paying customers

There is also an extra space at the bottom for ‘lost customers’: the customers that don’t make the conversion step (grey).

How to use the Pirate Metrics Canvas

The canvas is best used with a small team. Fill in the canvas the first time with what you know or guesstimates, and keep updating it during the innovation journey with real data to see how you might improve.

Fill in what you know

The first pass of the canvas is to fill in what you know already.

Step 1. Potential Customers

Find out how many Potential Customers your market has (roughly). Use demographics or for instance Facebook’s audience targeting system to get an idea of the maximum number of people you might reach. Stick these numbers on post-its in the potential customers area. It may be helpful to define customer segments when they have different preferred channels.

Next, find out how many of these potential customers you can already reach. What are the channels you already have (or would like to have) to reach potential customers? How many people interact with these channels? List them as well. This is probably a lower total amount.

Think of channels such as:

- Your existing website and social media
- Partner websites and social media
- Advertising
- Blogs
- …

Step 2. Acquisition

Define what ‘acquisition’ means for your situation. When do you count a customer as acquired? When they end up on your landing page? When they are on your email list? When they open the App Store page? Or something else?

For different channels, list conversion rates. How many people in the channel can you direct towards your service? If you place an ad, how many people click on it? If you write a blog, how many people read it? Try to find the typical conversion rates from your data (if you have it), or find appropriate benchmarks online. Later you'll be able to figure out what the actual rates are.

Step 3. Acquired customers

Using the conversion rates for each channel, figure out how many of the potential customers made it to this stage. How many people make it to your landing page? The rest has left and ended up in the 'lost customers' box.

Step 4. Activation

Define what 'activation' means for your situation. When do you count a customer as active? When they download and install your beta? When they posted on your forum? It has to be an activity they perform at least once and that is specific to what you're selling.

How many people will become active users of your service? Find conversion numbers in your data or look for suitable benchmarks online. Later you'll be able to figure out what the actual rates are.

Retention, Referral, and Revenue

Here it becomes a bit more complex. Each Activated customer can end up in zero, one, two, or even all three of the boxes to the right.

Step 5A. Retention

This is the most obvious step to look at. How many active customers come back to use the service again. When they come back, that means the service has value for them. They are obviously happy about at least some aspects. Improving this rate means you are able to zoom in on how you deliver value to the customer, and how you make your service addictive (Check out the awesome book *Hooked* by Nir Eyal to read more about addictive services).

Retention is a really important metric, especially for subscription based models. Most successful businesses get upwards of 70% of their revenue from loyal customers, and improving retention (lowering churn) by even 5% can mean a difference of 95% in your revenue.

The longer a customer stays in a subscription model, the higher their CLTV will be, and therefore, relatively, the lower their CAC is.

Important metrics startups use that are influenced by retention are ‘daily active users’ and ‘churn’.

Again, try to get numbers from the data you have, or look at online benchmarks.

Step 5B. Referral

This is arguably the most important thing to look at if you want high growth and lower CAC. Dropbox famously influenced their referral rates by giving free storage for bringing in new users. Referrals mean free advertising. People referring your service give it a warm introduction to others: their referral is much more likely to be trusted. It also means that people need to be quite convinced of how much they like your service before they refer it.

As before, try to get numbers from the data you have, or look at online benchmarks. Also come up with an idea of how many people they will refer. When they take the decision to refer people, and especially when there is an incentive such as with Dropbox, they will likely refer more than one person.

Step 5C. Revenue

This is the bottom line. How many people end up paying for the service. If you have a freemium model, for instance, not every active user may be a paying customer.

Again, come up with a conversion rate.

Step 6A. Retained users

Calculate how many users will be retained based on the number of active users and the retention rate.

Step 6B. Referred users

How many people will be referred to your service? This is a function of the number of active users, the referral rate, and the number of people referred on average.

Step 6C. Paying customers

Calculate how many people will pay for the service.

Strengths and Weaknesses

Now that you have filled the canvas, take a step back. With the team, identify strengths and weaknesses. Where are you doing better or worse than your benchmarks? Where do you have missing information? Mark these out with coloured dots.

Then, go over each step. For each step:

1. Come up with 3 ways to influence the conversion rate (negatively or positively). This could be adding a different channel, sending a personal message, making something easier to understand, …
2. Come up with 3 reasons for a customer to convert to the next stage.
3. Come up with 3 reasons for a customer to become a lost customer. Try to focus on things you can influence.

Stick these on the board on post-its.

Note that in some cases it's great that people don't convert: if you already know a person is not going to be an active customer, for instance, it's better they don't pass the acquisition stage. The funnel is also a filter to find those customers that are the best fit for your service.

Define Experiments

Use the strong/weak spots and your calculated numbers to find likely candidates for improvement. Where can you influence the end result in the most effective way? That should be your focus for the near future.

Also, if you haven't done so, setup the necessary tooling to measure the different transitions and stages in your model.

Revisit

After a month or so, revisit your canvas, and update the numbers with real data. What have you learned? Where were your numbers spot on? Where were they way off? Try to see if you understand what is going on. And then iterate: find the next weak spot, and define a new experiment.

Thank you for reading the Experiment Cookbook!

The cookbook was created based on 10 years of experience in innovation and venture building in 2019. It has been downloaded over 5000 times since then. I hope it brings you what you have been looking for!

Cheers,
Erik

[END]