

the right ^{*}it

* Why So Many Ideas Fail
and How to Make Sure
Yours Succeed

Alberto Savoia

Care has been taken to confirm the accuracy of the information presented and to describe generally sound and accepted practices and techniques. However, neither the author nor the publisher shall be held liable or responsible to any person or entity with respect to any loss or incidental or consequential damages caused, or alleged to be caused, directly or indirectly by the information contained therein.

THE RIGHT IT. Copyright © 2019 by Alberto Savoia. All rights reserved. Printed in the United States of America. No part of this book may be used or reproduced in any manner whatsoever without written permission except in the case of brief quotations embodied in critical articles and reviews. For information, address HarperCollins Publishers, 195 Broadway, New York, NY 10007.

HarperCollins books may be purchased for educational, business, or sales promotional use. For information, please email the Special Markets Department at SPsales@harpercollins.com.

FIRST EDITION

Designed by Terry McGrath

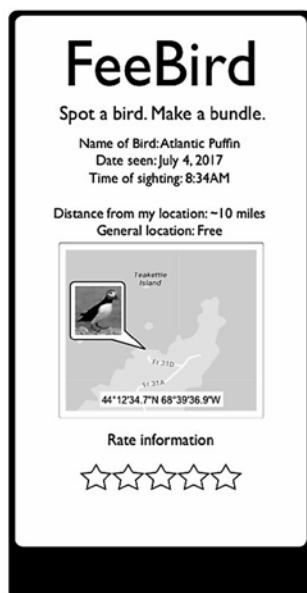
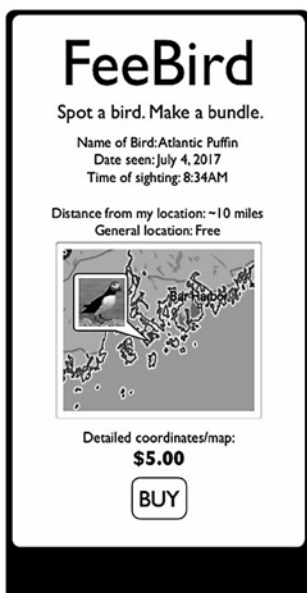
Library of Congress Cataloging-in-Publication Data is available upon request.

ISBN 978-0-06-288465-7

19 20 21 22 23 LSC 10 9 8 7 6 5 4 3 2 1

Below are mock-ups of two FeeBird screens I put together on Apple Keynote in about ten minutes. The first screen shows the general location of an interesting bird near you, along with an option to buy a precise map/location for a fee of \$5:

The second screen shows what happens if users decide to buy the detailed information: they get a detailed map, GPS coordinates, and a chance to rate the information.



The question is, after people get to see the Tortell-o-matic in action, how do I assign skin-in-the-game points to the responses I get from the market? That’s where a Skin-in-the-Game Caliper comes in handy. As an example, here’s the Skin-in-the-Game Caliper for my Tortell-o-matic:

Type of Evidence	Examples	Skin-in-the-Game Points
Opinion (expert or nonexpert)	“Great idea.” “Nobody will buy it.”	0
Encouragement or discouragement	“Go for it!” “Keep your day job.”	0
Throwaway or fake email address or phone number	bogusemail@spam.com, (123) 555-1212	0
Comments or likes on social media	“This idea sucks,” thumbs-up or thumbs-down, Like	0
Surveys, polls, interviews online or off	“How likely are you to buy on a scale of 1-5: ____.”	0
A validated email address with the explicit understanding that it will be used for product updates and information	“Give us your email to receive updates about the product: ____.”	1
A validated phone number with the explicit understanding that you will be called for product updates and information	“Give us your phone number so we can call you about our product: (____) ____-____.”	10
Time commitment	Come to a 30-minute product demonstration	30 (1pt./min.)
Cash deposit	Pay \$50 to be on the waiting list	50 (1pt./\$)
Placing an order	Pay \$250 to buy one of the first 10 units when available	250 (1pt./\$)

The TRI Meter

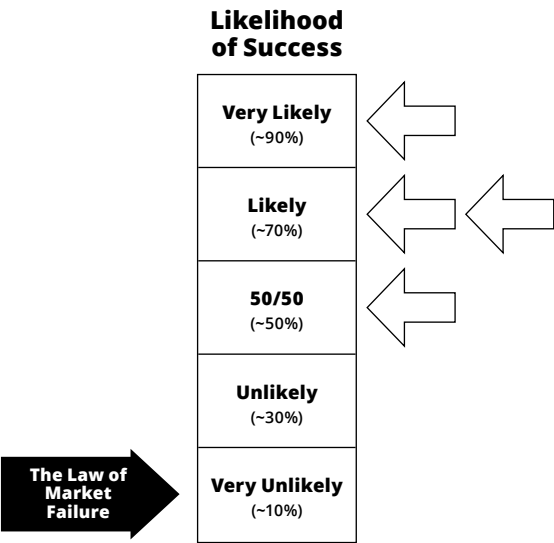
Collecting YODA with skin in the game to validate our Market Engagement Hypothesis is a necessary first step, but raw data by itself is not sufficient. In order to extract value out of data and use it to make rational and well-informed decisions, we need a way to interpret it, put it on a scale, compare it, and combine it with other relevant data.

The data from a cholesterol test, for example, is just a ratio of two values, the number of milligrams of cholesterol per deciliter of blood. Let's say you go for your annual physical exam, take a blood test, and learn that your total cholesterol is 300. That number by itself is not very meaningful, but when the doctor pulls out a chart and shows you that, statistically speaking, people with a cholesterol level of 300 are 4.5 times as likely to die of heart disease than those with a cholesterol level of 200, you may decide to lay off cheeseburgers for a while.

The Right It Meter, TRI Meter for short, is a visual analysis tool I developed to help you interpret the YODA you've collected as objectively as possible. More precisely, the TRI Meter is a gauge to help you estimate how likely it is that an idea will succeed in the market, but it's a gauge that is not too technical, complicated, or confusing—which can easily happen whenever probability and statistics are involved.

First let me show you what the TRI Meter looks like in action. Then I will explain how to use and interpret it. The image below shows a TRI Meter after four prototyping experiments (represented by the four white arrows on the right). As you can see, the TRI Meter scale is partitioned into five likelihood-of-success categories ranging from Very Unlikely (10% chance of success) to Very

Likely (90% chance of success), each representing the likelihood that your idea is The Right It.



Here's a guideline to help you map your answer on the TRI Meter:

If the data **significantly exceeds** what the hypothesis predicts, you point the arrow to **Very Likely**.

If the data **meets or slightly exceeds** what the hypothesis predicts, you point the arrow to **Likely**.

If the data **falls a bit short** of what the hypothesis predicts, you point the arrow to **Unlikely**.

If the data **falls really short** of what the hypothesis predicts, you point the arrow to **Very Unlikely**.

Finally, if the data is for some reason **ambiguous, potentially corrupted, or hard to interpret**, you point the arrow to **50/50** or optionally discard it. After all, even in science not all experiments produce clean and reliable data.

Example: Second-Day Sushi

All of this sounds more complicated than it actually is, so let me show you how to use the TRI Meter on our Second-Day Sushi example, which, unlike the fish itself, should still be fresh in your mind. First, we have to make sure that we have an XYZ Hypothesis that we can hypozoom into a set of xyz hypotheses.

If you recall, we had the following XYZ Hypothesis for Second-Day Sushi:

At least 20% of packaged-sushi eaters will try Second-Day Sushi if it's half the price of regular packaged sushi.

And from that, we hypozoomed to our first xyz hypothesis:

At least 20% of students buying packaged sushi at Coupa Café today at lunch will choose Second-Day Sushi if it's half the price of regular packaged sushi.

To test the first xyz hypothesis, we came up with a Relabel prototype: stick a label that says “Second-Day Sushi: 1/2 Off” on half of the boxes on display and count how many people choose to buy them. Let’s assume that 100 boxes of sushi are on display and that we label half of them (50 boxes) as Second-Day Sushi. The key piece of data we need to collect is the percentage of Second-Day Sushi boxes sold compared to the total number of sushi boxes sold. In other words, how many people who wanted sushi for lunch chose to buy Second-Day Sushi?

Let’s assume that during lunch, students bought a total of 40 boxes of prepackaged sushi. How many of those boxes were the relabeled Second-Day Sushi? Here are a few possible scenarios:

Outcome	Number of Boxes Sold out of 40	% of Total Boxes
A	0	0
B	2	5
C	6	15
D	8	20
E	16	40
F*	2	5
G**	30	75
<p>*On the day of the experiment, there was an article in <i>The Stanford Daily</i> about the risks of eating raw fish.</p> <p>**The lunch crowd included a group of 130 Japanese students visiting the campus.</p>		

Here’s how these results map on the TRI Meter:

Outcome A (0% of all boxes sold): Ask yourself, “If Second-Day Sushi is The Right It, how likely is it that we would sell 0

boxes out of forty?” Given that your first xyz hypothesis predicts 8 boxes and you sold 0, this is an easy call. The arrow from that result should point to Very Unlikely.

Outcome B (5% of all boxes sold): This result is not as discouraging as the previous one; after all, two people bought into the idea of slightly stale sushi, but it’s well below the predictions of 20% from our hypothesis. Unless we are ready to dramatically revise our business model and expectations (e.g., target a few daring souls who are *really* short on cash and hungry for sushi), outcome B should also point to Very Unlikely.

Outcome C (15% of all boxes sold): The data from this experiment provides evidence that some sort of viable market might exist, but that that market is not as big as we need it to be to make the business successful based on our existing hypothesis. For now, and unless we decide to adjust our business model and expectations accordingly, these data point to Unlikely.

Outcome D (20% of all boxes sold): This is at the bottom range of our hypothesized market, but it fully meets the minimum requirements for confirming the hypothesis. Yay! It deserves a Likely rating.

Outcome E (40% of all boxes sold): Whoa! This result blows our prediction out of the water. If we ask, “If Second-Day Sushi is The Right It, how likely is it that we would sell 16 boxes out of forty?” we can confidently answer Very Likely.

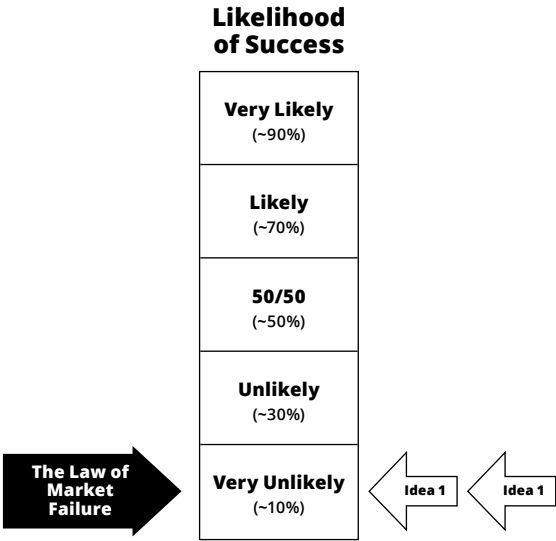
Outcome F (5% of all boxes sold): This is a discouraging result, but the data is questionable because, by an unlucky coincidence,

the school's daily paper had a front-page article on the risks of eating raw fish. Because of that, we either point the arrow to 50/50 (inconclusive) or discard it.

Outcome G (75% of all boxes sold): This is a fantastic result, but we have to be objective, so we cannot ignore the fact that on the day of the experiment an unusually high number of high-school students on a college tour from Japan visited the café. Perhaps these young students did not fully understand the implication of the Second-Day Sushi name, or perhaps they did not have a lot of money for lunch. Either way, since this was not a normal situation, we should probably dismiss this particular result. As much as we'd like to believe that our idea is the greatest ever, we have to be careful not to fool ourselves.

Round 1: Punched in the Face

Let’s begin with the most common scenario. Unless you get lucky, after a couple of experiments the TRI Meter for the first iteration of your idea (Idea 1) will look something like this:



If you are new at this, those first punches are going to surprise, hurt, and disorient you. But don’t let this kind of result demoralize or discourage you.

First of all, welcome to the club! What club? The very crowded club of people who thought that their idea was for sure—no doubt about it—The Right It, only to have their hopes and expectations mercilessly dashed by the Beast of Failure.

Second, think how much worse off you’d be if you had gone ahead with that idea without testing it. After investing months of work and lots of money to develop and market your product, you find out that your idea was The Wrong It all along—a knockout punch that sends you to the hospital. Fortunately, our thinking,

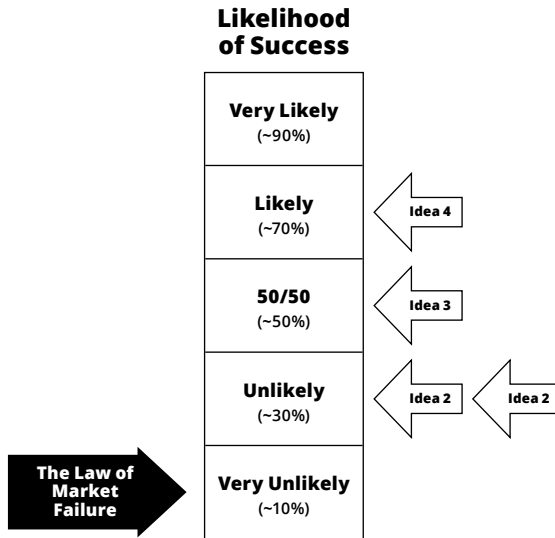
prototyping, and analysis tools can help you avoid that. A little pain now can save you tons of pain later. By learning quickly and cheaply that a particular idea is not likely to be successful, you will have plenty of time and resources left to modify your original idea or explore a new set of ideas—to go a few more rounds.

Based on this TRI Meter, we should concede that our beloved new product idea is most likely headed for failure in the market. Round 1 goes to the Beast of Failure. If you are really passionate about your new product, you may decide to get back into the ring and run a few more experiments with the same exact idea—just to be sure. But a more logical and less painful course of action would be to go back to the drawing board (or back to your corner, if you like the boxing metaphor) and use what you’ve learned from your experiment to tweak your idea.

Rounds 2–4: We Take Some, We Give Some

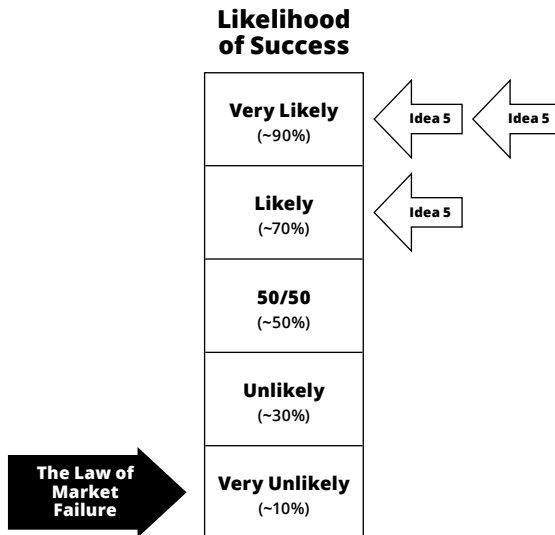
We make some tweaks to our original idea (Idea 1) and run some tests with each of the variations (Ideas 2, 3, and 4). When we map the results on the TRI Meter we get the “Likelihood of Success” shown at the top of page 170.

We still get punched quite a bit—especially with Idea 2—but not as hard as before. Our tweaked versions of the idea manage to stay out of the Very Unlikely zone, and we even manage to land a punch with the fourth version of our idea (Idea 4). That’s a very good sign—we are learning more about the market, tweaking accordingly, and moving closer to The Right It territory.



Round 5: We Land a Few Good Punches

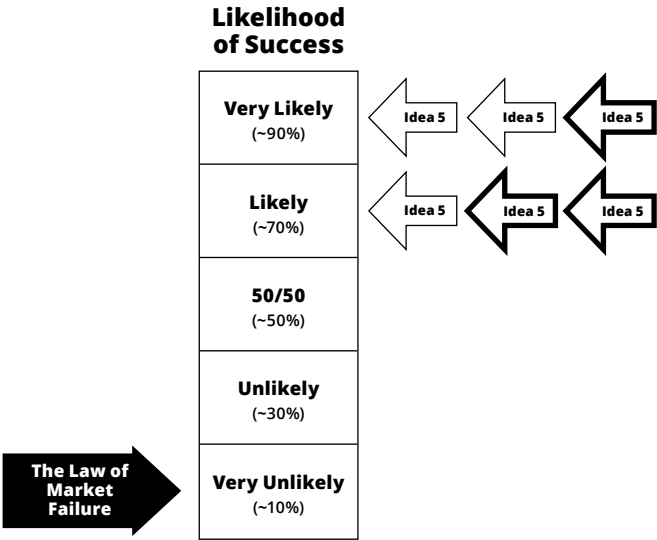
Using Idea 4 (the one that scored a Likely) as the starting point, we make a couple of additional tweaks and go back into the ring with Idea 5.



The arrows from our three experiments with the fifth version of our idea all point to Likely or Very Likely. This is great! Assuming that the experiments that produced those results were properly designed and run and that the data from each was interpreted fairly and objectively, there’s strong evidence that this idea might be The Right It. But that ominous black arrow at the bottom does its job and reminds us how rare it is for a new idea to succeed in the market. Are those three positive results enough to balance and counteract the Law of Market Failure?

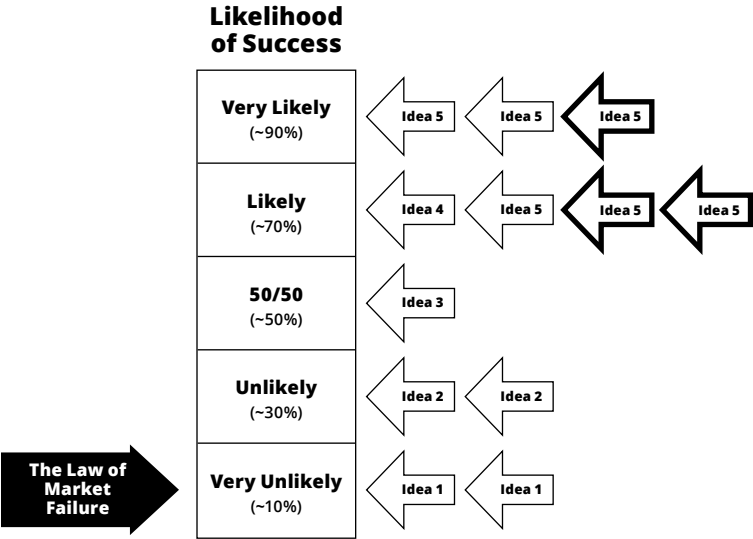
Developing this particular idea will require a major investment and commitment, and we want a higher degree of confidence before going ahead. So we decide to run three more experiments using Idea 5.

We map the new results on the TRI Meter alongside the first set of results (second set of results shown in **bold**), and we get the following:



All right! The new set of experiments on Idea 5 confirms the first set of results. This is great. We can't completely ignore the big black arrow—the market may still surprise us with an unexpected punch—but there's a good chance that the fifth iteration of our idea is The Right It.

To help you visualize the process, here's what our sequence of tweaks and experiments looks like if we chart them all (from Idea 1 to Idea 5) on a single TRI Meter:

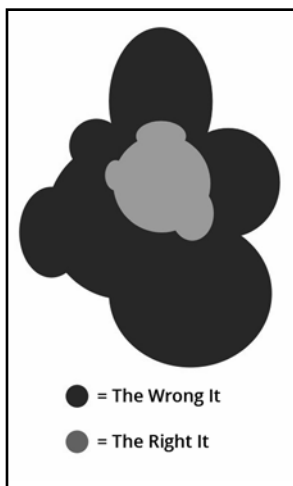


We ran a total of twelve prototyping experiments on five different ideas (or versions of a similar idea). That may sound like a lot of tweaking and experimenting, but with prototyping this would not have taken more than a couple of weeks—less time than what most teams would spend to write an OPD-based business plan.

As we wrap up our discussion of the TRI Meter, let me repeat that only arrows that represent actual data from carefully designed and personally conducted experiments are allowed. No

opinions and no OPD (market research done by other people, with other methods, at other times—you know the drill). Your arrows must consist of only YODA with skin in the game.

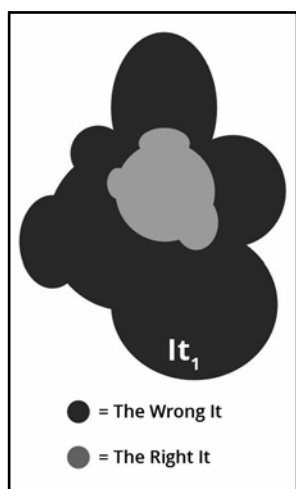
In the following image, the light gray area represents The Right It zone for a specific market opportunity, while the dark gray area represents The Wrong It zone.



The market opportunity is there, and it's real. But not all products that address it will be successful: they may be too expensive, too big, too complicated, come in the wrong color, have the wrong name, and so forth. The market can be very, very fussy and fastidious. If you don't come up with a product combination it likes (i.e., The Right It), it will reject your idea—even if, in many other respects, your product does a good job of

addressing the problem or opportunity. If you are really committed, or interested, or passionate about a specific market problem or opportunity, stick with that market, but tweak and experiment with variations on the original idea.

Let's assume that you've run several prototyping experiments

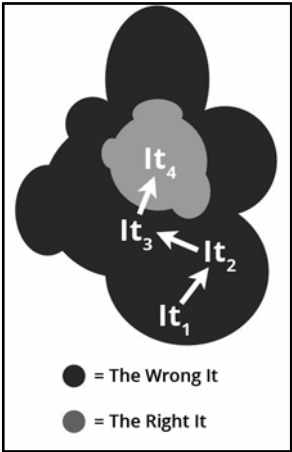


with your initial idea (represented as It_1 in the next image), and the data shows conclusively that your idea—as it currently stands—is The Wrong It.

You are disappointed—and that's understandable. But in the process of testing that idea, you have discovered interesting facts about your target market. Perhaps you learned that one of your key Thoughtland-based assumptions was dead wrong (e.g., most people *don't* think that \$8 is too expensive for packaged sushi). Or you

observed that 80% of packaged-sushi shoppers carefully inspect the label to check the “packaged on” time stamp—and put the box back if that time indicates that the box is more than a day old. With each experiment, you gained valuable YODA you can use to inform and guide your next step.

Even if the original idea for Second-Day Sushi proves to be The Wrong It, it is quite plausible that somewhere in the Thoughtland neighborhood from which Second-Day Sushi originated dwells an idea and business model for budget packaged sushi that is The Right It. Try testing a weekly subscription service and tweak the name and slogan to suggest convenience rather than a lack of freshness: “Sushi2You: The convenient and affordable sushi subscription service.”

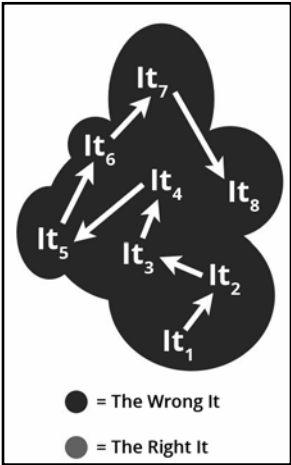


the image below shows, The Right It for the budget-sushi idea does not exist; all variations on that idea are destined to fail, because too many people associate cheap sushi with bad sushi—and bad sushi with all sorts of nasty consequences.

And if the market does not respond to that particular tweak, explore and pretotype a few more tweaks (e.g., “Groupie Sushi: Order sushi as a group and save”) until you find a combination that is The Right It.

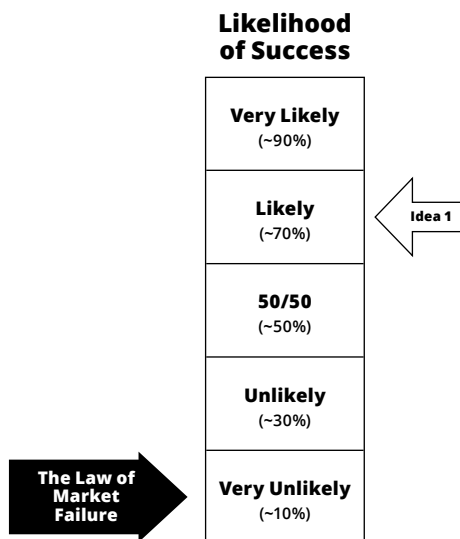
But what if, even after a bunch of tweaks, you still haven’t found The Right It?

At that point you should consider the very real possibility that,



Analyzing and Iterating

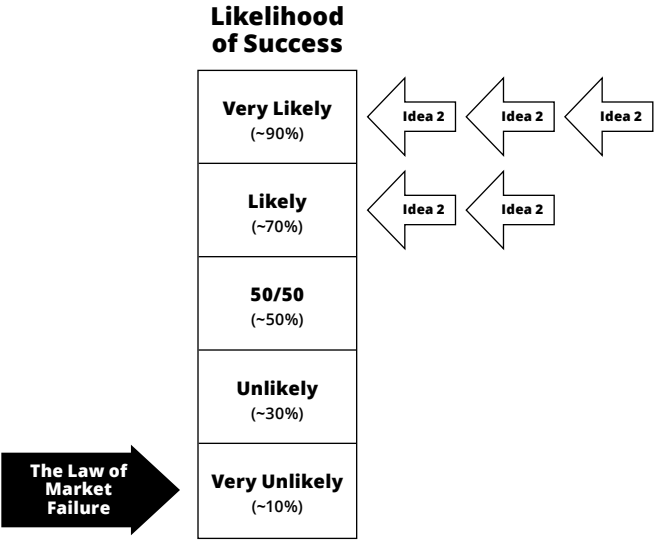
Hypothesis xyz1 predicted a 40% response, and our test returned a very healthy 62%. This result exceeds our expectations and indicates a strong level of interest. Normally, a number this good (i.e., substantially better than our estimate) is an indication that an idea is Very Likely to succeed. But since we did not mention the hefty \$3,000 price tag in our email and many Google workers may have assumed that the BusU classes would be free (or be reimbursed by Google), I decide to be conservative in my interpretation of the result. Instead of Very Likely, I score it as Likely.



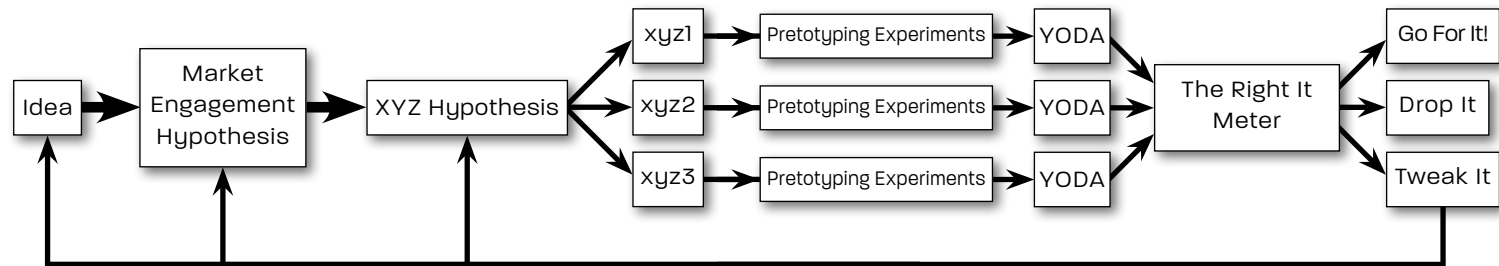
I could have been even more conservative and scored the experimental result as 50/50 or tossed it away as bad data. But getting a greater than 60% response to any idea in any market happens so rarely that I decided to take it as evidence of strong market interest, and “If there’s a market, there’s a way.” The next set of experiments will determine whether I am right or wrong.

We have an encouraging first result, but the big black arrow on the TRI Meter is doing its job. It's keeping us grounded in reality. It reminds us that most new ideas will fail in the market and that we need to have several more successful experiments—a preponderance of positive evidence—to balance out the Law of Market Failure.

We update the XYZ and xyz hypotheses and run a few more prototyping experiments to validate this new model. In the process we collect even more YODA and learn additional valuable lessons (e.g., we can earn an additional \$60 per student by selling coffee and snacks on the bus). Here’s what the TRI Meter for the BusU business model for \$300 classes (Idea 2) looks like after five experiments:



It looks as if the new version of BusU has a good chance of being The Right It, wouldn’t you agree?



glossary

Beast of Failure: An imaginary merciless and insatiable creature that devours most new product ideas and bites and slashes those who pursue those ideas without first validating them.

Data beats opinions: A key rule you need to internalize and practice with no exceptions if you want to improve your odds for market success. Don't base your product decisions on opinions. Base them on market data, and not just any old data or other people's data (see OPD)—but your own (see YODA).

DTD (Distance to Data): A metric to help you quantify and minimize how far you have to go to collect your market data.

\$TD (Dollars to Data): A metric to help you quantify and minimize how much it will cost to collect your market data.

expert opinions: *See* opinions.

Failure is not an option: An inspiring but wrong and often harmful phrase and belief suitable for Hollywood movies but ill suited for most business ventures. Use it only if you are writing dialogue for an action flick with Arnold Schwarzenegger.

false negative: An idea that is mocked and knocked in Thoughtland but, when competently executed and brought to market, turns out to be a success.

false positive: An idea that sounds great in Thoughtland, is competently executed and brought to market, and fails miserably. Most market failures are false positives.

HTD (Hours to Data): A metric to help you quantify and minimize how long it will take to collect your market data.

hypozooming: The process of taking an XYZ Hypothesis and “zooming” it in with regard to scope, space, and time to derive a set of smaller related xyz hypotheses that can be tested locally, quickly, and inexpensively. For example, the xyz hypothesis for Second-Day Sushi: “At least 20% of students buying packaged sushi at Coupa Café today at lunch will choose Second-Day Sushi if it’s half the price of regular packaged sushi.” The principle is that if the XYZ Hypothesis is true, then the easily testable hypozoomed xyz hypotheses derived from it will reflect that.

If there’s a market, there’s a way: An important reminder that if there’s enough market interest in an idea for a product, you (or someone else) will usually be able to find a way around whatever engineering, financial, legal, or other obstacles currently stand between the idea and its realization. If an idea is The Right It, one way or another it will usually find its way to market.

If there’s no market, there’s no way: If there is no market interest in an idea, no amount of brilliant design, amazing engineering, superior reliability, or marketing fireworks will help the idea succeed.

If we build it, they will come: An overly optimistic, ill-founded sentiment not suitable for business ventures, because they *won't* come, unless you are building The Right It. However, by transposing two words and adding a question mark, you get the critical question you should ask before building anything: “If we build it, will they come?” That’s the key question this book helps you answer.

Law of Market Failure: “Most new products will fail in the market, even if competently executed”; the stubborn hard fact that most new product ideas are destined to fail and that competent execution of an idea cannot save it from failure.

market failure: When the actual market result from an investment in a new product is less than or the opposite of the expected result.

market success: When the actual market result from an investment in a new product meets or exceeds the expected result.

MEH (Market Engagement Hypothesis): A high-level description that combines the fundamental premise behind a new product idea with a vision of how the target market will engage with it. For example, the MEH for Second-Day Sushi: “A lot of people who want to eat healthy and who like sushi can’t afford to eat it regularly, because sushi tends to be quite expensive. If we can find a way to make sushi as affordable as other fast food, a lot of fast-food patrons will choose sushi over less healthy alternatives.”

OPD (Other People’s Data): Market data collected by other people at other times, in other places, with other methods, using other filters, and for other purposes. Technically, OPD is data, but

since it's not *your* data, it can be as dangerous and as misleading as opinions. OPD is no substitute for YODA (*see* YODA), and given that OPD is neither necessary nor sufficient for evaluating your ideas, it's best to not waste time seeking it.

opinions: Subjective, biased, and often baseless judgments about an idea's prospects for success. In the quest for The Right It, opinions are worse than useless—they are downright dangerous and misleading.

pretotype: A specific artifact or technique used in prototyping. Some pretotypes are the Mechanical Turk, the Pinocchio, the Fake Door, the Facade, the YouTube, the One-Night Stand, the Infiltrator, and the Relabel.

prototyping: Using a group of tools and techniques to collect fresh, reliable, and relevant market data (*see* YODA) about an idea for a new product as quickly and as inexpensively as possible. The goal of prototyping is to help you *make sure that you are building The Right It before you build It right*.

Say it with numbers: A reminder to quantify whenever possible. A number, even if it's just an educated guess at the time, is more informative and useful than some fuzzy, vague term. For example, instead of "Our widget will be inexpensive," say, "Our widget will cost \$10," or "Our widget will cost 40% less than the competition's."

skin in the game: Something *of value* given to you by the market as evidence of interest in your idea. The simplest form of skin in the game is money (e.g., a paid preorder or a deposit), but it could also be someone's time, information, reputation, and so on.

Skin-in-the-Game Caliper: A tool to help you quantify and calibrate skin in the game, necessary because not all skin in the game is created equal. For example, a \$1,000 preorder for a product should count more than a \$100 deposit for the same product, and a commitment to attend a one-hour presentation should count more than, say, an email address.

Thoughtland: An imaginary place where ideas for new products hang around collecting solicited and unsolicited opinions. Like visits to Las Vegas, visits to Thoughtland are best kept short: most of what happens in Thoughtland should stay in Thoughtland.

The Right It: An idea for a new product (or service, company, initiative, etc.) that, *if* competently executed, will succeed in the market. The Right It is like Kryptonite to the Beast of Failure.

The Wrong It: An idea for a new product (or service, company, initiative, etc.) that, *even if* competently executed, will fail in the market. The Wrong It is like catnip to the Beast of Failure.

TRI Meter (The Right It Meter): A tool for mapping and visualizing the results of pretotyping experiments to help determine how likely it is that an idea is The Right It.

XYZ Hypothesis: What you get when you apply “say it with numbers” to the Market Engagement Hypothesis. The basic form for the XYZ Hypothesis is: “At least X% of Y will Z,” where X% represents a percentage of your target market, Y, and Z represents how that percentage of the market will engage with your new product idea. For example, the XYZ Hypothesis for Second-Day Sushi: “At least 20% of packaged-sushi eaters will try Second-Day Sushi if it’s half the price of regular packaged sushi.”

xyz hypothesis: A small, specific, easily and quickly testable hypothesis that is derived from, and consistent with, a broader XYZ Hypothesis. For example, a possible xyz hypothesis for Second-Day Sushi is: “At least 20% of students buying packaged sushi at Coupa Café today at lunch will choose Second-Day Sushi if it’s half the price of regular packaged sushi.” The process of going from a broad XYZ Hypothesis to one or more xyz hypotheses is called *hypozooming*.

YODA (Your Own Data): Data about *your own* product idea, collected firsthand by *you and your own team*, by running experiments *you* designed to validate *your own* market hypotheses. To qualify as such, YODA must come with skin in the game (*see skin in the game*). Unlike OPD, YODA is both necessary and sufficient for evaluating your idea.