

Teaching Lean Startup Principles: An Empirical Study on Assumption Prioritization

Matthias Gutbrod and Jürgen Münch

Faculty of Informatics, Reutlingen University, Reutlingen, Germany
[matthias.gutbrod|juergen.muench@reutlingen-university.de]

Abstract. Creating new business models, products or services is challenging in fast-changing unpredictable environments. Often, product teams need to make many assumptions (e.g., assumptions about future demands) that might not be true. These assumptions impose risks to the success and these risks need to be mitigated early. One of the principles of the Lean Startup approach is to identify and prioritize the riskiest assumptions in order to validate them as early as possible. This helps to avoid wasting effort and time. In the literature there are several different methods for identifying and prioritizing the riskiest assumptions reported. However, only little research exists about the practical application of these methods in practice and how to teach them. In this paper, we present and empirically analyze a workshop format that we have developed for teaching the prioritization of Lean Startup assumptions. We aim at raising the awareness for assumption thinking among the participants and teach them through group work how to prioritize assumptions. The results of the analysis of a multitude of conducted workshops show that the applied method did lead to reasonable results and accompanying learning effects. In addition, the participants got aware of assumption thinking and liked learning in a practical way.

Keywords: Risk Prioritization, Riskiest Assumptions, Teaching, Lean Startup.

1 Introduction

Kevin Systrom had an idea for a location sharing app where users could “check-in”, called Burbn. The programming of the iPhone app took him a few months. In Burbn users could check-in with friends who are hanging around, get points and take and post pictures. The app had many features and was therefore complicated to use. The app was unsuccessful, but Kevin Systrom started together with Mike Krieger to analyze what the customers really were doing with their app. They found that the original assumption that users will use a “check-in” feature was wrong. This could have been validated before the full implementation of the feature. The two observed that users were basically only posting photos. Together, they decided to get rid of all of the app functionalities except for sharing and liking photos. They spent months creating and experimenting with prototypes in order to validate risky assumptions. In the end, they built an app called Scotch. Scotch was slow and full of bugs. Nevertheless, they doubled down on

the insight that sharing photos in a frictionless way is important for users. In the next version, they focused on a super easy to use app where the users only need three clicks to upload a photo. They called the app Instagram and launched it in October 2010 [10].

Creating new products or services is quite challenging because there is a high risk of creating something that nobody wants [4]. More than half of all product ideas do not work. Typical reasons are that customers are not excited about a product or that a product is too difficult or time-consuming to use. Sometimes, there are problems with the business viability due to legal, financial or business constraints [11]. Many assumptions are made during product development that come from team members or superiors. Product teams, for instance, try to take the customers perspective and they imagine that customer use a product in a specific way. When they observe real customer behavior afterwards, they are often surprised that customers behave quite differently. Due to Gladstone [9], "it is often the unexpected way that a customer uses a product, that highlights it is true potential".

In order to raise the odds of success of product and service development it is important to identify the important assumptions that need to be true for success. These assumptions need to be validated as early as possible. An important task is to identify these assumptions. But how to find them? Where are they documented? Usually, all relevant aspects of a product or business idea are documented in canvas models such as the Business Model Canvas [3] or the Lean Canvas [4]. At the beginning, these canvas models are full of untested assumptions. Therefore, canvas models can be seen as a good starting point for identifying assumptions.

Every entry in a business model is an assumption until we have proven that the assumption is right. Assumption testing is an essential activity [13]. However, product development is limited by time and other resources so that not every assumption can be tested. This is the reason why we should first identify which assumptions are the riskiest ones and test them first. Ries states that "Lean Startup is designed to operate in [...] situations where we face [...] extreme uncertainty..." [1]. Ries calls the riskiest assumptions "Leap-of-Faith Assumptions" (LOFA). They can be seen as claims in a business plan that will have the greatest impact on its success or failure. Very often, LOFAs focus in the beginning around the problem and the customer segment. Testing these assumptions is quite difficult as customers "often think they know what they want, but it turns out that they are wrong" [1]. Careful validation techniques, e.g. through customer development interviews, is necessary to validate those assumptions.

There are many methods dealing with risk prioritization, but there are only little research and practical experience on teaching them. In this paper, we describe a workshop format that guides participants on how to prioritize assumptions of an example business models or business ideas.

The rest of this paper is organized as follows: Section 2 presents related work. Section 3 defines the research approach and the research questions. In Section 4 we present the results followed by Section 5 with a discussion and lessons learned. Section 6 summarizes the paper and outlines future research.

2 Related Work

There exist several different approaches for prioritizing assumptions with respect to risks. In this section, we describe some of the popular methods. The methods have several differences: some are, for instance, using matrixes with dimensions, others are based on quantitative risk calculations, and some methods recommended specific sequences in which assumptions can be tested to reduce risks.

The first matrix approach is the Prioritization Matrix by J. Gothelf and J. Seiden [5]. They use the two dimensions “known to unknown” and “low risk to high risk” in order to classify and compare different assumptions. The second matrix approach is the Prioritizing Leap-of-Faith Assumptions (LOFA) matrix described in the book “The Startup Way” by Eric Ries [1]. Ries also uses two dimensions. The first one is the “time of impact” which describes when the assumption will have an impact. The second one is the “magnitude of impact” which describes how big the impact is on the business model if the assumptions are false [1].

The approach by J. Fjeld consists of a calculation with three parameters: severity, probability and cost. After all the parameters are calculated for every assumption, they can be ranked [6]. Ash Maurya divided risk into three different categories: product risk, customer risk and market risk. He recommends to prioritizing the assumptions based on the stage of your product [4]. The book “Value Proposition Design” by Alexander Osterwalder et al. [7] includes also a description of a simple prioritization method: a long line from the bottom “less critical to survival” to the top “critical to survival” is used for prioritization. The hypotheses from a business model can be pinned along this line and ranked in order [7]. Another method is described in the book “Disciplined Entrepreneurship” by Bill Aulet [12]: he recommends making a list of all the areas in which logical conclusions have been made, such as conclusions about producers, consumer and development. Laura Klein presents a method in her book “Building Better Products” [14] that is based on a risk identification grid with two separate factors. The first factor describes how likely it is that an assumption is true and the second factor is how bad the outcome will be if it is not true. For further details we refer to a previously published more comprehensive analysis of risk prioritization methods that has been conducted by the authors of this article [8].

3 Research Approach

In order to teach how to prioritize Lean Startup assumptions and to raise assumption thinking we created a workshop format that uses one of the aforementioned prioritization methods. For the workshop we selected as initial method Eric Ries’ method from “The Startup Way” which is based on sorting assumptions along the dimensions “time to impact” and “magnitude of impact”. The assumptions are mapped onto a matrix. All assumptions in the top right quarter, which have a high magnitude and a near impact can be seen as LOFAs. They should be tested with experiments as early in the product development process.

In the workshops, for the specific task of risk prioritization we wanted to teach the participants how they can easily classify assumptions into a risk matrix and identify the riskiest ones. We prepared original Airbnb assumptions, so the participants did not have to make their own assumptions initially. There are 22 assumptions in total that the participants worked with. All participants were divided into groups of 3 to 6 participants. At the beginning of the workshop we gave a presentation motivating the relevance of the topic and explaining what assumptions are in the context of Lean Startup. Additionally, we showed them some examples how startups identified and tested their assumptions. Directly before the risk prioritization task, the participants got a short introduction about Airbnb.

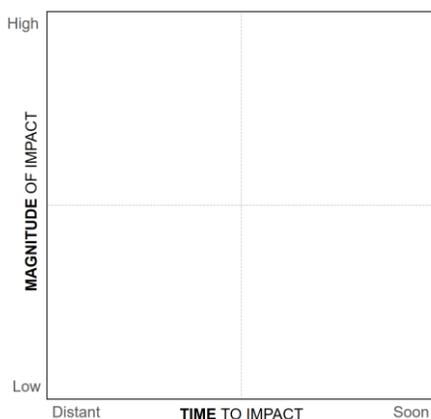


Fig. 1. - Leap-of-Faith Assumptions Matrix

(© Eric Ries - The Startup Way p.93)

After that, we showed them the task and explained the two axes of the Leap-of-Faith Assumption matrix. Additionally, we explained to the participants that the riskiest assumptions go in the upper right corner of the matrix where the distance from each axis is the greatest. During the task, the participants worked completely alone without any help. Each group got a poster with the assumption matrix. The assumptions were already written on prepared sticky notes. The participants could look at all the assumptions and potential relationships between different assumptions and decide where to put them in the matrix. Each group had 20-25 minutes for this task. After the task, the results were photographed with a camera and discussed. Each group was allowed to present the three to four riskiest assumptions they identified.

With this approach the following research questions should be answered:

- RQ1: Did the teams identify the riskiest assumptions?
- RQ2: Which assumptions are particularly correct/wrong categorized?

4 Execution and Analysis

A total of 6 workshops was carried out with 19 teams and in consequence a total result of 19 matrixes. Immediately after each workshop task, we captured the results in pictures. All pictures were copied to a digital folder and then individually printed on pages. After printing, we measured the length of each axis in the bottom left quarter. We used the results to scale the manually measured points to full-scale. Then, using a ruler, we manually measured the distance from each axis to the sticky notes with the assumptions. All the information from the measurements were recorded in Excel. Subsequently, we measured the same quarter on the bottom left of the original poster, took the value and set it in relation to the previous manual measurements. With this value, we scaled up all the manually measured points to the original size. The gained data was used for the analysis of the team results. In total, there were 397 sticky notes with 22 different assumptions.

4.1 Findings

Figure 1 shows the distribution of the assumptions in the prioritization matrix. The groups classified the assumptions into two different dimensions. The first dimension was “Time to Impact” and the second was “Magnitude of Impact”. Each dot represents one of the 22 assumptions. The position in the matrix represents the average positioning of the 19 groups for each assumption with respect to these two dimensions.

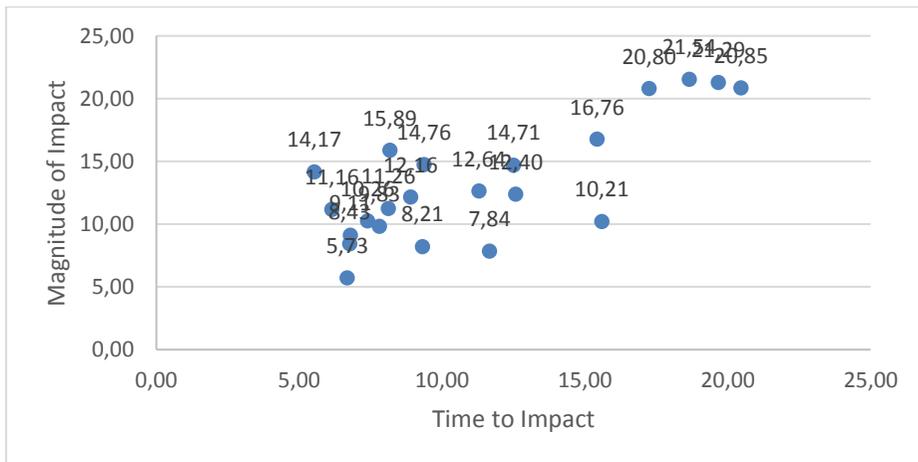


Fig. 2. - Distribution of all assumptions in the prioritization matrix

Table 1 presents the individual assumptions together with the average prioritization result for each individual assumption along the two dimensions. In addition, the standard deviation indicates the degree of agreement between the different teams with respect to the prioritization.

Table 1. – Prioritization results for each assumption and standard deviation

Id	Items	Assumption	TIME Ø	STD- ABW	MAGNI- TUDE Ø	STD- ABW
1	18	Travelers are willing to rent from strangers (no hotels)	20,46	4,53	20,85	2,73
2	18	Homeowners will allow strangers to live with them for a short time	19,67	3,34	21,29	2,29
3	18	AIRBNB is legal	18,65	4,82	21,54	2,63
4	19	There are enough travelers who accept the offer	17,25	5,92	20,80	2,31
5	19	Travelers seek alternative experiences to hotels	15,42	4,53	16,76	5,13
6	19	Adding pictures to descriptions results in a higher booking rate	12,50	6,8	14,71	5,3
7	16	Hotels are perceived as impersonal	15,59	3,84	10,21	5,89
8	16	Travelers find hotels too expensive	12,58	6,36	12,40	6,35
9	19	Displaying apartments on a map will result in more bookings	9,37	4,76	14,76	5,46
10	8	An indication of the ratings below the description of the apartments increases the likelihood of a booking	8,17	4,59	15,89	4,4
11	19	The cleanliness of the apartments is the most important thing for the travelers	11,29	5,92	12,64	4,37
12	19	It is very important for the travelers to check-in to the rented apartment at any time	8,90	4,88	12,16	5,6
13	19	Design A of the search page leads to more bookings like Design B	5,53	5,52	14,17	6,77
14	18	Travelers do not want to clean up after their stay	11,66	6,91	7,84	4
15	19	Through a Google Adwords campaign we get more bookings	8,12	5,58	11,26	5,24
16	19	The faster the search results appear, the higher the probability of a booking	7,39	5,11	10,26	4,53
17	19	Profile pictures of the travelers increase the likelihood that the landlords will accept the booking	7,81	5,12	9,83	5,39
18	19	Travelers like to pay for cleaning	9,32	5,15	8,21	4,06
19	19	Pictures by professional photographers lead to a higher booking rate	6,14	4,38	11,16	5,61
20	19	By sending vouchers to customers we get more recurring bookings	6,79	3,75	9,11	5,09
21	19	In the result of the apartment search it is better to see more than 5 hits	6,76	6,29	8,43	4,7
22	19	The price selection is easier with a sliding scale	6,68	4,49	5,73	3,2

Overall, the common agreement on the “Magnitude of Impact” dimension was greater than on the “Time to Impact” dimension.

In the “Time to Impact” dimension, the groups identified the following three assumptions as the most important: "Travelers are willing to rent from strangers (no hotels)", "Homeowners will allow strangers to live with them for a short time" and "AIRBNB is legal". The smallest standard deviation of the dimension was $\sigma = 3,34$ with the assumption "Homeowners want to allow strangers to live with them for a short time". The largest standard deviation was the assumption: "Travelers do not want to clean up after their stay" with a value of $\sigma = 6.91$.

With respect to the “Magnitude of Impact” dimension, the groups identified the following three assumptions as the most important: "AIRBNB is legal", "Homeowners want to allow strangers to live with them for a short time" and "Travelers are willing to rent from strangers (no hotels)". The smallest standard deviation of the dimension was $\sigma = 2,29$ with the assumption "Homeowners will allow strangers to live with them for a short time". The largest standard deviation was the assumption "Design A of the search page leads to more bookings like Design B" with a value of $\sigma = 6.77$.

One of the interesting results was that all groups from every workshop had always independently identified the same three riskiest assumptions.

4.2 Threats to Validity

In this section, we critically discuss our study results regarding internal and external threats to validity:

How comparable is the business model to other business models? In the selected example, we are dealing with a platform business model. Although, many business

models fall into this category of business models, other archetypes of business models such as direct business models exist and this might impact the results.

The criteria “Time to Impact” and “Magnitude of Impact” were chosen as prioritization criteria. Other criteria can also play a role, such as the effort involved in testing. We chose this risk matrix as an initial approach to prioritize assumptions because it is proposed by Eric Ries popularized the Lean Startup approach.

Can the method we described be used outside of a workshop? We have tested the method and overseen its use in workshops. That does not necessarily mean the method works online.

Did the teams understand the prepared assumptions and were they clearly formulated? The teams explained to each other how they understood the prepared assumptions and ended up with a common vision. Two times, teams asked for the meaning of an assumption because they did not understand it correctly.

Is our evaluation correct? Were the results well photographed and are they usable? We tried to photograph the group results from a direct position as best as possible. The results were printed out on A4 paper and measured manually with a ruler. Some small inaccuracies remain. Firstly, the sticky notes had no exact reference point so we had to choose them freely. Secondly, some of the sticky notes were overlapping, making it difficult to set the reference points. The calculation was carried out with the help of Excel and was additionally controlled by another researcher.

What kind of prior knowledge did the participants have to bring along for the described part of the workshop? The participants needed to know the Airbnb business model to understand the assumptions. At the beginning of the workshop, we first asked whether they knew the business model or not. We then briefly explained what Airbnb does and we placed an Airbnb info sheet on each of the group tables. How were the appropriate Airbnb assumptions selected? The assumptions were made by the scientists using various sources of literature. Together, we selected the assumptions for the workshop and the selection was subjective.

Are the assumptions simply unfounded? Attention was paid to ensure that the assumptions were understandable, therefore other scientists were shown the assumptions and questioned if their meaning was clear.

Are the assumptions too simple and do not represent real assumptions? We extracted the assumptions from real Airbnb reports and books so we believe that the assumptions can be thoroughly tested.

In order to generalize the results, further research with more workshops and training is necessary.

5 Lessons Learned and Discussion

Overall, the workshops were well received and the participants had no major problems in conducting the tasks. The following lessons learned could be identified:

Providing an example case with a set of predefined assumptions seems to be an easy and efficient way to teach the concepts. This worked very well for the participants. The participants could get immediately involved in the task of risk prioritization and did not

have to spend much time for coming up with their own assumptions. Additionally, the prepared assumptions had the advantage that the participants had no personal feelings about them so that they could view the assumptions more objectively.

The groups were randomly created, so that they typically consisted of participants with different backgrounds. The participants learned that there were different opinions on where to put the assumptions in the matrix and needed to come up with an agreement. The participants in each group were able to get a common understanding. Usually, the groups needed 20-25 minutes to map all assumptions on the matrix. In one workshop the group size was bigger, i.e., 5-6 persons per group. In this case the mapping took around 35 minutes.

If a team struggled with the classification of an assumption on the time dimension, it helped to give them a hint: "Think about the following: Which assumption needs to be successfully tested first". This helped the participants to better arrange their assumptions on the time dimension.

The two dimensions were quickly understood and there were rarely questions about the dimensions.

Working with the assumptions was fun for the participants and they gained a newfound awareness that identifying, understanding, prioritizing, and validating important assumptions is a highly relevant activity. After the task, some participants recognized that working with assumptions and testing the riskiest ones can also create significant effort.

After the workshops, the participants often asked if they could take the risk matrix poster and the used material home. This indicates that the participants have understood the importance of risk prioritization and that they are interested in applying this method to their very own business and product ideas.

6 Outlook

We plan to make improvements to the workshop materials so that a simpler and even more accurate analysis of the results is possible. The sticky notes will be provided with a reference point and a number in order to better measure the exact position in the matrix and to better support the analysis. We are also planning to conduct short qualitative interviews after the group work in order to complement the analysis. Further workshops are planned to increase the significance of the results.

Another research avenue we are currently discussing is to develop software-based simulators so that participants can learn prioritization online and/or by using more than one prepared business scenario.

References

1. Ries E.: *The Startup Way*. , Portfolio Penguin, 2017, pp. 89-93, p. 99
2. Cooper B., Vlaskovits, P.: *The Lean Entrepreneur*, Second Edition, Wiley, 2016, pp. 127
3. Osterwalder A., Pigneur Y., *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*, Wiley, 2010

4. Maurya A.: Running Lean, 2012, Second Edition, O'Reilly
5. Gothelf J., Seiden J.: Lean Ux Applying Lean Principles to Improve User Experience, 2013, First Edition, O'Reilly, pp. 22
6. J. Fjeld: How to Test Your Assumptions, MIT Sloan Management Review, Winter 2018, pp. 89-90
7. Osterwalder A., Pigneur Y., Bernada G.,Smith A. :Value Proposition Design, Wiley, 2014, pp. 203
8. M. Gutbrod, J. Münch and M. Tichy : “The Business Experiments Navigator (BEN)”, IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC) 2018 vol 18
9. B. Gladstone: “<https://www.codegent.com/blog/2014/3/assumption-testing>”, Codegent Blog 2014
10. K. Sawyer: Zig Zag: The Surprising Path to greater Creativity, Wiley, 2013, pp 22
11. M. Cagan: Inspired - How to create Tech Products Customer Love, Wiley, Second Edition, 2018, pp 111
12. B. Aulet: Disciplined Entrepreneurship, Wiley, First Edition, 2013, pp. 219
13. J. Liedtka and T.Ogilvie: Designing for Growth, Wiley, First Edition, 2013, pp.131
14. L. Klein: Building Better Products, Wiley, First Edition, 2016, pp. 245-247