

# How Do Software Startups Approach Experimentation? Empirical Results from a Qualitative Interview Study

Matthias Gutbrod<sup>1</sup>, Jürgen Münch<sup>1</sup>, Matthias Tichy<sup>2</sup>

<sup>1</sup> Faculty of Informatics, Reutlingen University, Alteburgstraße 150, 72762 Reutlingen, Germany

<sup>2</sup> Institute of Software Engineering and Programming Languages, Ulm University, Germany  
[matthias.gutbrod|juergen.muench@reutlingen-university.de, matthias.tichy@uni-ulm.de]

**Abstract.** Software startups often make assumptions about the problems and customers they are addressing as well as the market and the solutions they are developing. Testing the right assumptions early is a means to mitigate risks. Approaches such as Lean Startup foster this kind of testing by applying experimentation as part of a constant build-measure-learn feedback loop. The existing research on how software startups approach experimentation is very limited. In this study, we focus on understanding how software startups approach experimentation and identify challenges and advantages with respect to conducting experiments. To achieve this, we conducted a qualitative interview study. The initial results show that startups often spent a disproportionate amount of time focusing on creating solutions without testing critical assumptions. Main reasons are the lack of awareness, that these assumptions can be tested early and a lack of knowledge and support on how to identify, prioritize and test these assumptions. However, startups understand the need for testing risky assumptions and are open to conducting experiments.

**Keywords:** Experimentation, Experiment, Software Startups, Lean Startup, Minimum Viable Product.

## 1 Introduction

Drew Houston, the co-founder and CEO of Dropbox, got his idea for developing a file-sharing tool on a long bus ride to New York when he wanted to work but could not because he had forgotten his USB stick [9]. Developing a file sharing tool such as Dropbox requires significant resources such as time, effort, and money. The founders of Dropbox wanted to avoid waking up after years of development to see that nobody wants their product. Therefore, they decided to run a small experiment in order to test the most critical assumption, i.e., to test if most people have the problem of file synchronization and would give the product a try. They built a three-minute demo video and uploaded it to Hacker news together with a call to action to join the waiting list for the private beta program. The video “drove hundreds of thousands of people to the website”, Houston reported after the experiment and continued, “our beta waiting list went from 5,000 people to 75,000 people literally overnight. It totally blew us away.” With this experiment, the founders validated the most critical assumption that there was

real interest for their product [8]. The development of new products and services in startups typically faces many uncertainties. For all these uncertainties, startups need to make assumptions based on their current knowledge. Startups need to transform the riskiest assumptions into testable hypotheses and test them early, in order to avoid proceeding based on the wrong assumptions, which could have high-risk business impacts. Different techniques for testing assumptions can be used. These techniques are all fundamentally based on scientific experimentation principles. One approach that focuses on experimentation is the Lean Startup approach which has been popularized by Eric Ries [8]. It fosters hypothesis testing by applying experimentation as part of a constant build-measure-learn feedback loop. The build-measure-learn feedback loop can be seen as a motor that should always run and produce learning data about critical assumptions. There is only limited research about how startups approach experimentation. In this study, we focus on understanding how especially software startups approach experimentation and identify challenges and advantages with respect to conducting experiments.

The rest of this paper is organized as follows: Section 2 presents related work. Section 3 defines the research questions and describes how the study was executed. In Section 4 we present the findings followed by Section 5 with a discussion. Section 6 summarizes the paper and outline future research.

## 2 Related Work

The scientific literature offers several frameworks for experiment-driven product development based on empirical findings. Bosch, for instance, proposes a framework for building products as innovation experiment systems [1]. Fagerholm et al. present building blocks for a continuous experimentation system and an infrastructure [2].

Several studies exist that focus on dealing with uncertainties. Nquyen-Duc et al. emphasize the role of prototyping for experimenting with business ideas. In a study, they identified factors influencing a prototype-centric learning loop. One important finding from their study is that it is necessary to align the prototyping approach with the learning goals [7]. Another study by Nguyen-Duc and Abrahamsson explored the role of minimum viable products (MVPs) in early stage startups. Among other findings, they found that MVPs play an important role in bridging knowledge gaps between the entrepreneur team and stakeholders [6]. These knowledge gaps could be seen as critical assumptions that need to be tested. Other studies focus on the challenges and benefits of organizing and conducting business-oriented experiments. Bosch et al. conducted a study that was aimed at typical challenges when finding a product idea worth scaling. They found that only very few companies “worked with continuously validating product concepts with customers to try to identify problems before building a full solution.” Based on their findings they identified several key areas where support is needed. One of these key areas is the validation of product ideas [4]. Hassi and Tuulenmäki focused in a study on how to organize the innovation practices for experimentation. One result was that the design and organization of innovation activities has a major impact on the ability to unlock the full potential of experimentation [3]. In a study, Lindgren and

Münch explored the state of the practice of experimentation in the software industry. The study found that experimentation is rarely done systematically and continuously in practice [5]. The study described in this paper focusses on understanding how software startups approach experimentation and identifies challenges and advantages with respect to conducting experiments in startups.

### 3 Research Approach

In practice, the actual situations in which startups use experiments are not clear. And if they are using experimentation, how do they conduct the experiments? To get a better understanding we defined the following Research Questions:

RQ1: *How do startups use experimentation in practice?*

RQ2: *What challenges do startups have with conducting experiments?*

RQ3: *What benefits do startups see with respect to experimentation?*

Given the exploratory nature of our study and the "how" research question, we decided to use a qualitative interview study in the context of a multiple case study approach [10]. All selected cases were software startups in their early phases before the product market fit with different products. With the term software startups, we refer to human institutions searching for scalable, repeatable, and profitable business models in order to create new software-based products or services, in the context of extreme uncertainty and unpredictable dynamic technology markets. The data collection method was semi-structured interviews with open-ended questions. The interview duration time was between 15-75 minutes. At the beginning and end of each interview, we included a "warm-up" and "cool down" question. Before every interview, the participants got a declaration of consent form to sign, which included ethical guidelines and the data privacy protection. All interview participants have the role of CIO or CTO in their company, of which they are also the co-founder. Each interview was recorded and transcribed for detailed analysis. The data analysis followed procedures suggested by Yin [10]. The interviews were conducted with German startups companies by the primary researcher from March to June 2017. Table 1 presents a short profile overview.

**Table 1.** Profiles of the software startups cases

Software startup	Business domain	Founded	Number of Founders	Current product development phase
Case 1	Software as a Service	2015	3	Functional product with a few large customers.
Case 2	Online marketplace	2016	4	Functional product
Case 3	Job portal	2017	3	Prototype
Case 4	Software as a Service	2016	2	Functional prototype

## 4 Results

**Case 1:** This startup provides a web-based toolbox for designing web applications that can be connected to the external and internal services of companies. The solution aims at helping companies to digitize their actual processes. It can be used without any programming skills and provides many reusable modules from a toolbox.

The startup conducts several kinds of experiments. It primarily conducts landing page tests. The main goal of these tests is to increase the conversion rate and transform more visitors into prospective buyers. The CEO gave an example in the interview: *“On our landing page we use A/B testing for positioning call to action buttons in order to increase the conversion rates.”* He added that the duration of such experiments depends on how the team members vote, and is on average 2-3 weeks.

The startup also does implicit experimentation with some kinds of minimum viable products (MVPs) and calls customers regularly to test assumptions. The startup also exhibits at trade fairs and uses these opportunities to do small experiments.

One of the main challenges the startup faces is to drive enough visitors to its landing pages so that the results from the A/B tests are significant. The CEO mentioned: *“At the beginning we were happy if we had two visitors on our page”*. He also mentioned technical challenges with respect to A/B testing: *“We change the content on the website with JavaScript and this can lead to conflicts with the actual JavaScript code on the page.”*

**Case 2:** This startup focuses on developing a web-based market place for vintage cars. The marketplace consists of an online platform for selling and buying. In addition to the online platform, the startup provides a hybrid mobile app. The founder and CEO of this startup explained that the startup is currently not conducting experiments in a systematic way. However, when developing new features or making changes to the current system, he feels a need to better understand customers: *“Sometimes you ask yourself after the 10th start of the application: are the customers happy? Do they have new ideas for improvements?”* Currently, new product ideas mainly stem from the startup itself: *“We develop the ideas in our heads. There are only a few which come from customers.”* The startup visits fairs such as trade fairs, and uses these opportunities to test assumptions about products and features through interviews with potential visitors and investors. According to the CEO, the main challenge that hinders the conduction of experiments is the lack of resources. He mentioned: *“We are 2-8 people in the startup and there is no time to do experiments. However, in the future we will do more in this direction.”*

The CEO of this startup had already founded another startup before, that focused on the creation of a speech recognition app. Based on the vision of this startup and ideas for features that were proposed by customers, the startup created MVPs and observed, how new features resonated with users. The startup acquired users for these MVP tests from their friends, through their networks and through Facebook groups. The CEO mentioned that some testers were very engaged: *“We had a super contact from Munich, which tested our software very deeply and wrote us 50-60 pages reports with weakness or strengths from our product. All for free.”* The CEO also mentioned that he used some

testers outside of the target customer segment and gained interesting insights from these testers. The startup also conducted customer interviews and A/B tests. The duration of experimentation was from 2 hours to 2 weeks.

Another strategy that the CEO applied at his former startup was to partner with influencers: "We had some meetings with influencer with about 4-5 million followers in social media channels. These influencers started to use their audience to test our product assumptions. In this way we got important insights before the implementation".

The CEO summarized his thinking about testing product ideas as the following: "You should first find out how customers like your product before you go to market. We recognized several times that the customers understood our product ideas in another way than we expected. It happened that 9 from 10 people said: 'go that direction not the other'. In case that ideas from customers were tested successfully and got implemented, the customer identified themselves with the product and told others: 'Look, they did this because of me'. In this case, the customers are feeling that they are the co-creator and this is gold".

**Case 3:** This startup has the vision to create a web-based online platform to connect good job applicants to the right companies. With an intelligent psychological matching system, it aims at finding the perfect match between a job offer and an applicant. Additionally, the startup wants to help people to better understand different education and career choices in order to select those that fit their preferences and abilities. This aims at opening the spectrum of options with respect to education and career and to help people avoid choosing the wrong path.

The CTO of the startup explained that the first experiments were telephone interviews: *"It costs a lot of time and kilometers to drive to different companies and ask their hr manager questions. On the phone, we could find the responsible person faster"*. The questions were prepared in the following manner: *"We made a short brainstorming session and fixed 16 questions, from which some were around our idea. But in the first calls we found out that the interview partners did not have enough time. Therefore, we reduced the questions to eight."* The experiment duration was one week.

A challenge with conducting interviews was making cold calls. In addition, the CTO mentioned in the interview: "I have a little problem to motivate myself make calls for a whole week." A second challenge was to convince the companies to talk with the startup and answer the questions. The CTO already had experience with contacting customers from his last job as a support person. However, the other founders had no experience. In addition, he explained that *"you should know how to talk to a person on a phone and how to convince them to answer our questions."*

Actually, the first kind of customer-related study that the startup conducted was a market research study with data from statistica.com. Based on this study, the startup concluded that the assumed problem exists. In consequence, the startup started to develop a solution partly based on wrong assumptions. The CTO explained: *"I wish we could turn back time and do the interview study first and get its results. We lost 1,5 years."*

**Case 4:** This startup has the vision to support persons at a gym with a training software. The startup developed a combination of hardware and software as a solution. This startup started directly with the development of the product. The CEO explained: *“We built our product because the core was clear. He mentioned how the startup comes up with ideas for new features and improvements: “At first, we try to take the customer’s perspective. This is always the fastest way. We also use our friends and they give us feedback.” He adds that friends are good testers, but they have a weakness: “It is hard to get a honest feedback, if we do interviews with our friends. We use our own question catalogue.” We asked the CEO why he did not create a landing page for the product. He answered that he considers the product as an innovation: “We want to keep the range of the people under control who are knowing our product.”*

Table 2 gives a summary of the experiment types and the perceived challenges and benefits found in the studied software startups.

**Table 2.** Experimentation Summary overview

Conducted Experiment Types	Perceived Challenges with Experimentation	Perceived Benefits from Experimentation
Interviews Trade show testing Landing page A/B testing MVP testing Testing with influencers	Getting enough subjects for experiments Fear of contacting customers Fear of making cold calls Technical challenges with setting up an infrastructure for experimentation Lack of skills for conducting customer interviews Lack of resources/staff for experimentation Lack of motivation to conduct experiments Fear that people steal the startup idea	Early feedback Better understanding of customer’s needs, priorities, and behaviors Better prioritization of development activities Avoidance of unnecessary development efforts Early testing of market demand

## 5 Discussion

The studied software startups mainly address these uncertainties by getting feedback about product ideas from potential customers or friends. They basically use qualitative techniques such as unstructured or structured interviews to test ideas, or more light-weight approaches such as testing ideas at trade fairs.

Another approach that some of the studied startups are using is to create initial versions of products or features, expose them to existing or potential customers, and observe customer reactions in an unsystematic way. To some extent, this is similar to experimentation with an MVP, although a clear hypothesis or a learning goal is usually not explicitly stated.

Some of the studied startups use specific types of experiments for specific purposes, such as A/B testing in order to optimize a web site design, or landing pages for testing value propositions and generating leads. Our study shows significant challenges that

startups are facing with respect to experimentation. Examples are problems with getting enough subjects, lack of motivation, and technical challenges. A major challenge identified in this study is the lack of knowledge about experimentation techniques and startup tactics, such as tactics for finding experimental subjects or showing potentially secret product ideas to others. Although software startups already apply techniques such as qualitative customer interviews that are well suited for problem exploration and solution testing, they do not have sufficient knowledge on how to use these techniques in a way that creates valuable insights.

The lack of knowledge about experimentation techniques and startup tactics might be a reason for the identified fears that hinder startups to conduct experiments. Another challenge that the study reveals is a lack of support for startups with respect to conducting experiments.

Such support might comprise training (e.g., how to conduct customer interviews), technical expertise (i.e., how to implement A/B tests), and infrastructure for experimentation (i.e., infrastructure for observing customer behavior). The studied startups lack the awareness that risk mitigation needs to be done early and is crucial for survival. In addition, it is not well known that risks can be mitigated early and systematically through experiments. Indicators for this lack of awareness are that 1) all studied software startups focused on testing ideas and none of them systematically validated the problem and the customer segments (usually, both should be done before validating solutions), 2) the statement by the CEO from case three that they have difficulties in motivating themselves and thus infects the other founders to conduct experiments negative. The lack of awareness that risks can be mitigated early and systematically through experiments is probably the most important challenge as it is an “unknown unknown” that significantly threatens the success of a startup. The integration of startups in accelerator programs, and a more wide-spread entrepreneurship education might help to overcome this challenge. It should be mentioned, that the studied software startups have shown a widely positive attitude towards experimentation. The startups understood some of the major benefits of experimentation, especially with respect to avoiding solution risks. In addition, customer centricity was seen as a valuable means towards developing successful products and services.

The study has several shortcomings with respect to validity. The first validity threat is related to generalization. Qualitative case studies are a suitable approach to understand how software startups use experimentation in the real world. To get more generalizable results, more case studies and quantitative studies should be conducted. The second validity threat is the construct validity which especially addresses misunderstandings between the researchers and participants. Before each interview, the participant got a short introduction about the goals of the study. Furthermore, the use of the interview guide enables asking clear questions. Another validity threat is related to the degree of knowledge about the history of the startup. To diminish this risk, we only interviewed founders or cofounders who generally have the best knowledge of their startup process.

## 6 Conclusions

The studied software startups apply some ad hoc validation of solution-related testing, the majority of the studied startups are not aware of the importance of risk mitigation and have a lack of knowledge on how to do this. Simply speaking, the studied startups do not systematically identify the riskiest assumptions, they do not have sufficient knowledge on how to describe such assumptions in a testable way, and they do not know how to test assumptions efficiently by means of experiments. Finally, they do not have sufficient support for conducting experiments in terms of training, technical competence, and infrastructure. Despite this, they are open to learn and improve their capabilities in order to increase the odds of their success.

We are planning to further investigate on how startups approach experimentation and on how to better support them in doing so. Further studies need to be conducted in order to increase the generalizability of the results.

**Acknowledgements.** We wish to thank all participants for their time and contributions.

## References

1. Bosch, J. (ed.): Building Products as Innovation Experiment Systems. ICSOB. Springer-Verlag Berlin Heidelberg (2012)
2. Fagerholm, F., Sanchez Guinea, A., Mäenpää, H., Münch, J.: The RIGHT model for Continuous Experimentation. *Journal of Systems and Software* **123**, 292–305 (2017). doi: 10.1016/j.jss.2016.03.034
3. Hassi, L., Tuulenmäki, A. (eds.): Experimentation-driven approach to innovation: developing novel offerings through experiments. ISPIIM Conference, Manchester, Manchester (2012)
4. Jan Bosch, Helena Holmström Olsson, Jens Björk, and Jens Ljungblad: LNBIP 167 - The Early Stage Software Startup Development Model: A Framework for Operationalizing Lean Principles in Software Startups, 15 (2013)
5. Lindgren, E., Münch, J.: Raising the odds of success. The current state of experimentation in product development. *Information and Software Technology* **77**, 80–91 (2016). doi: 10.1016/j.infsof.2016.04.008
6. Nguyen-Duc, A., Abrahamsson, P. (eds.): Minimum Viable Product or Multiple Facet Product? The Role of MVP in Software Startups. XP2016 (2016)
7. Nguyen-Duc, A., Wang, X., Abrahamsson, P. (eds.): What Influences the Speed of Prototyping? An Empirical Investigation of Twenty Software Startups. XP2017 (2017)
8. Ries, E.: The lean startup. How Constant Innovation Creates Radically Successful Businesses / Eric Ries. The lean series. Portfolio Penguin, London (2011)
9. Victoria Barret: <https://www.forbes.com/sites/victoriabarret/2011/10/18/drop-box-the-inside-story-of-techs-hottest-startup> (2011)
10. Yin, R.K.: Case study research. Design and methods / Robert K. Yin. SAGE, Los Angeles (2014)